CHARTS AND GRAPHS

AN INTRODUCTION TO GRAPHIC METHODS
IN THE
CONTROL AND ANALYSIS OF STATISTICS

By

KARL G. KARSTEN, B. A. (Oxon.)
CONSULTING STATISTICIAN

INTRODUCTION BY CARL SNYDER
CHIEF STATISTICIAN OF THE FEDERAL RESERVE BANK OF NEW YORK

"It may be laid down as almost a fundamental principle that the statistician who is to be successful in business must cultivate the graphic methods."—Leonard Ayers.

New York
PRENTICE-HALL, INC
1925

First printing, October, 1923.
Second printing, January, 1925.

PRINTED IN THE UNITED STATES OF AMERICA

To

E. D. K. and E. C. K.

In inadequate acknowledgment.

PREFACE

In its general structure, this book follows a philosophic, and not an encyclopedic, arrangement. It therefore incidentally supports the author's theory of a system of natural evolution of charts, in accordance with which all chart-forms fall into line with simple origins and clear channels of growth. In the light of this theory, there is no baffling heterogeneity, no confusion of purposes or principles, in all the immense multitude of existing graphic forms. On the contrary, that multitude resolves itself into a consistent, organic body of simple root-forms and logical combinations and developments. Not only can we allocate each form to its proper place in such a system, but we can often discover gaps in the system, and bring to light forms which, while not yet in use, have reason to be. As examples of such experience, the following more or less original methods may be mentioned.

The author is indebted to a host of friends and associates, and therefore cannot, in a wide sense, claim originality for any of the methods described. But, in so far as no plain trail leads from some of them to any individuals, and, in so far as their first use is believed to be made either in this book or in earlier work of the author, he may take some small responsibility for the summary-chart, the double-probabilities paper, the population-map, and the collapsible bead-map. The theory of the silhouette-bar curves, and the argument for the reversal of axes of ogives fall into the same class, as does the use of square-root paper for economic data, and the use of the names "amount-of-change" and "rate-of-change." The same statement holds true regarding the compounded-average-seasonal method. Needless to say, these are, almost all, inevitable results of the application of the theory that chart-forms are naturally and logically evolved, one from another.

The greatest contribution to chart-making, from any single source, is the Gantt Progress Chart. This chart is, unquestionably, the most powerful graphic device for business and for all

executive and managerial purposes. While the description has been rather full, as given herein, it is by no means complete; and the Gantt charting methods, in all their co-ordinated ramifications, constitute an independent system of accounting and of executive control, which goes far beyond the proper field of this book. The present volume must, therefore, be supplemented by another, Mr. Wallace Clark's "The Gantt Charts," to get the full benefits of the method. Mr. Clark's achievements in industrial engineering and the promotion of managerial efficiency are ample recommendation for his book. And the chart which has been so unqualifiedly praised and adopted by Mr. Fred J. Miller, former President of the American Society of Mechanical Engineers, by Mr. Walter N. Polakov, a leading authority on power engineering, and by European experts, needs no endorsement from statisticians.

Inadequate mention has been made in the text, of the work of Professor William F. Ogburn ("Social Change") on the geometric trend of human culture and civilization, which has gone far to influence the present writer in his presentation of the law of organic growth as the great *raison-d'etre* of rate-of-change curves. Professor Ogburn's careful and keen pioneer work in this field will have an increasing effect upon economic thought for a long time in the future.

In a different field, the work of Mr. Carl Snyder should be referred to in any discussion of chart methods as it has set a high standard in statistical research, and has, by graphic interpretation, given to abstruse economics a vital and practical bearing upon business and commercial welfare. He has been a leader in bringing mathematical skill, economic research, and business problems together. The student of chart-making can do no better than to study the methods used in the charts appearing in the *Monthly Review* of the New York Federal Reserve Bank, from which, as will be seen, we have drawn heavily for illustrative material.

Fewer, but as excellent, are the charts appearing in the *Bulletin* of the Cleveland Trust Company, prepared under the direction of Dr. Leonard P. Ayers. These charts, and Mr. Snyder's, are models. The charts of the Harvard Bureau of Economics, though of a single type, are always powerful and well-made. The charts in the *Monthly Survey of Current Business*, published by the Department of Commerce, are also well-drawn. Indeed, the use of good charts is steadily

increasing. We have seen few books so well illustrated with excellent charts as Dr. Ayers' "The War with Germany," or Mr. Joseph E. Pogue's "Economics of Petroleum."

Others to whom acknowledgment should be made, not alone for contributions to this particular volume, but for important contributions to the growth of a sound and efficient charting-practice, are Mr. John Wenzel and Mr. Arthur R. Burnet, both of whom, with the author, earlier enjoyed the privilege of working with that pioneer in the field, Mr. Willard C. Brinton. To many other economists and former associates, among whom may be mentioned Professor Robert E. Hale, Professor Paul Douglas, Mr. Stuart Chase, Mr. Paul Brissenden, Dr. Fred R. Macauly, Mr. John Scoville, and Mr. Richard Webster, the author is indebted in innumerable ways. The courteous permission of authors of books and articles in the same field, to borrow illustrations from their works, is appreciated, and the attempt has been made invariably to credit the sources of such illustrations as they appear in the text. It is a pleasure and a duty to recommend such important books as those of Lipka, Peddle, Running, Haskell, and Brinton; also the statistical treatises of Yule, Bowley, Secrist, King, and Kelley.

A word may be said as to the style of the text. It is a quaint and curious folk-way of the academic world that a technical account is worthy of respect directly in so far as it can *not* be understood. This hoary tradition is not limited to college walls—the rocky road to business, until recently, has rested on the self-same supposititious secrecy, and the paths of all professions lead to inner circles that guard, as best they can, the knowledge and the standards of their work. When such precautions make for better craftsmanship, they are most heartily to be endorsed. But, when they merely further selfish ends, they are a plague and pestilence, and those who practice them, only that their own minute monopolies of craft may be entrenched, come, sooner or later, into the class of parasites, retarding the growth of their profession.

Having confessed so little patience with the doctrine of the incomprehensible *per se*, we have naturally sought to empty the entire bag of tricks, and to tell the whole story of the chart in the simplest words that we command. Our belief has been that it is a lesser sin to be too easily understood than never understood at all. But at the same time, we have sought to make the story full and complete. If any of

our readers find charts which do not fall into place in this account, but would appear to have been omitted, we beg that they will freely advise and assist us to include them. It is, moreover, probable that, in spite of vigilant revision, many errors have crept in; we hope that readers who detect them will courteously co-operate by sending corrections, suggestions, and criticisms.

Chart-making is an art which all can practice. But there will always be a world of difference between the charts of amateurs and those of master-statisticians. Perhaps the day is not far off when, from the latter class, will come a group collectively intent on keeping up the standards, not the secrecy, of graphs and all statistics. The need for some criterion, high, but not too high to be effective, has been already felt, and efforts to establish safe statistical standards are on foot. It is our understanding that we may shortly look for sets of standard texts and examinations, from a committee of the American Statistical Association, under the chairmanship of Mr. Malcolm C. Rorty. Such steps will be warmly welcomed in the profession. The task is to set good standards and to make them public property in good plain everyday English. And as a contribution to the protection of the calibre of business statistics through the medium of graphic presentation, this book is offered.

Karl G. Karsten

New York City, September, 1923

TABLE OF CONTENTS

BOOK I. SIMPLE CHARTS

PART I. NON-MATHEMATICAL CHARTS

PART II. AMOUNT-OF-CHANGE ANALYSIS

BOOK II. ADVANCED CHARTS

PART III. RATE-OF-CHANGE ANALYSIS

PART IV. SPECIAL ANALYSES

PART V. CALCULATING CHARTS

Part VI. Two- and Three-Dimension Data

Part VII. Conclusion

Appendices

Bibliography

Indices

LIST OF ILLUSTRATIONS

By Chart-method

LIST OF ILLUSTRATIONS

By Subject-matter

INTRODUCTION

In and since the War the use and development of charts has been almost phenomenal—so large, indeed, that at least one able economist who is interested in such things thinks that we as a country have gone chart-mad. But this development has not been confined to this country, and it has a very solid basis in practical utility. There is little question that the chart represents a genuine saving in time and in mental effort.

In this it does not differ from the ordinary map. Suppose the mariner, the shipping clerk, or the school boy had to locate a given point on the earth with a statement, let us say, that it was two thousand miles southwest from London, twelve hundred miles south of New York, eight hundred miles north of Rio de Janeiro, and so on. All of this information might be useful and even, for certain purposes, necessary. It is, so to speak, the statistical data of the question. But it yields no *picture*. A map or a globe gives us this mental picture almost in a flash. And that is precisely the use and service of a chart. Let us take an example:

Within the last few months from this writing, the newspapers have been filled from day to day with reports of this or that industry making a "new high record." The figures give the idea of a prodigious boom, and, as we have so sadly learned to know, practically every boom is followed by a crash. So the wise man will shake his head at these "new high records," and sagely observe that "it cannot possibly last."

Well, in most industries with which we are acquainted, such new high records are the normal and usual thing, and the absence of them the abnormal. In other words, practically every industry, just like the population of the country, has a fairly steady rate of growth, and so, with sharp interruptions that come at more or less irregular intervals, it is the normal and characteristic thing that they should make these new high records. Naturally, such high records should at least not be regarded in the light of sensational news.

Let us take our old friend pig iron as an instance. We have monthly records of pig iron production running back for forty years. In twenty-four of those forty years *some* month of those years has made a "new high record" in pig iron production, that is, in 60 per cent of the cases.

Furthermore, these new peaks of production tend to run in sequences of four, five, and six years. So if we see an estimate that pig iron production for this year, let us say, will "break all records," we know that this is a rather foolish way of putting it, that it is simply the fairly normal thing and what we might reasonably expect in the absence of any powerfully disturbing causes like a world war or a profound depression in trade.

Now all this information you may laboriously dig out of the actual figures if you like, but you can get it all in a quarter or maybe a tenth of the time if it is spread out in chart form. Like the point on the map, all these relations there stand out vividly and almost instantly.

But it is not alone the economy of time and effort that is involved. The great thing, often, is that the chart will flash the thing not merely to the eye but to the mind; I mean that the picture gives you the idea of making the computation, and even that there is such a thing as a normal rate of growth, as in pig iron production. Lacking the picture, we might have little to prompt us to make the investigation or suggest even a hypothesis.

I know there are those to whom this easy method of mental traveling is not attractive, and even, perchance, a little irritating. Nothing else could explain, for example, why it is that our mathematicians should often go through long and laborious calculations in an endeavor to find out whether any close correlations exist between two sets of data, or whether a periodogram is going to fit a given set of figures sufficiently to make it the basis for a forecast, when there is a far quicker route. While recognizing to the full extent the value which these methods may have in competent hands, it is still literally true that thousands upon thousands of calculations of every kind and description have been made as to these degrees of correlation and all their like, involving hundreds and even thousands of hours of needless and useless work, when a near approximation in ninety per cent of the cases could generally have been obtained with a log chart in much less than an hour.

The typical mathematical bent of mind seems to luxuriate in difficulties, long calculations, and complicated formulae. The simple, swift, and direct seems to be foreign to its nature.

In our work at the Bank, we have had much reason to study attentively these normal rates of growth. It is quite astonishing to find how characteristic they are of the different industries, and different lines of trade, and even such things as the growth of bank deposits, money in circulation, and numerous other fluctuations of the modern economic world. They are so characteristic, in fact, that very often a log chart, with the figure for the average rate of growth in, let us say, the last twenty years, will suffice to identify the subject of the picture without further label.

But this idea of the persistence of growth, as a kind of a characteristic inertia in the different industries and trades, is certainly foreign to our present ideas about business or the thought of many economists. There are as yet few of our business men or industrialists, for example, who are now willing to believe that one can make a fairly good guess as to, say, the average production of pig iron, or the average railway traffic, or the average postal receipts for the years of 1930–33. It is almost certain that few industries or few enterprises are now planned with any long look into the future.

There are very notable exceptions, like the American Telephone and Telegraph Company and others that might be mentioned, where the work of development is planned out for years ahead. For most men, even in our large industrial enterprises, these are pretty much matters of rule of thumb or of year-to-year pressure. If it were not so, we should scarcely have such violent ups and downs of production and trade, the booms and slumps that bring such demoralization to industry and to profits, and so much needless suffering among the wage-earning population.

Some day we shall find a way around such stupidity, and it is my own belief that the most accessible avenue is through the grouping of the available data into interesting and well conceived charts. They are the most reliable and most stimulating instruments of education that we possess.

So I think it has been a worthy service that Mr. Karsten has performed in writing such an encyclopedic and exhaustive work upon the subject. The time is right for it, and it should be highly useful. I do not mean to suggest by this that the

mere making of charts is the whole story, any more than the possession of a fine hammer and a chisel makes a good carpenter. But it is certain that, without good tools, the best of artisans is badly handicapped, and I believe this is equally true of the business man and the director of large enterprises. He cannot but be going somewhat blindly if he does not have at his right hand, maps and charts of his whole work, extending years into the future, so that he may plan and anticipate in a truly prescient way.

The rest of the story is that such scientific recording and projecting into the future makes of business and industrial enterprise a kind of romance in reality. Even the most interesting of occupations gets to be a kind of humdrum routine, if we have no long look ahead. Nothing stimulates the imagination more than a well constructed excursion into the future. And in business enterprises this is almost impossible without the intelligent use of charts.

But there is more. So prodigious have our industrial activities as a nation become, so varied and so diversified, that it is given to few men, even the ablest, nowadays, to maintain any accurate and adequate idea of current business trends and developments, and carry on their own work at the same time. So I believe that soon our successful captain of industry, like the captain on the great ocean liner, will have always at his elbow a trained navigator or business pilot, who will supply him with the material wherewith to study his course and make his plans, and who will tell him at any given moment just where he is at! And such a navigator will find his most useful tool to be a first-hand working knowledge of the different forms of charts which this book describes.

CARL SNYDER.

New York, 1923.

BOOK I. SIMPLE CHARTS

PART I. NON-MATHEMATICAL CHARTS

Chapter I

MAPS AND DIAGRAMS

It is probable that the original diagrammatician lived many centuries ago, and it is not impossible that he was a cartographer. A search for him would lead us back to the days of "Captain Kidd" legend, when, judging by some records, the word "chart" invariably connoted a faded sketch of a lone island, dead trees, and buried treasure. It would lead us back to that intrepid explorer, Marco Polo, whose revisions of geography upset his contemporaries; back to the Arabs, whose excellent charts of the skies played so large a part in their nocturnal travels over the desert; and back to the Phœnicians, who doubtless kept strange maps to guide them about the Mediterranean shores and perhaps to warn them of dangerously shrewd villages where the bargaining was not profitable. We could not stop at the Egyptians, four thousand years ago, whose floor-plans of the pyramids have recently yielded up to us their secrets, nor at the Chinese whose six-thousand-year-old maps of the heavens have confirmed modern astronomical calculations of star movements. We should be carried back to prehistoric man, at least sixty thousand years ago, some of whose drawings have been identified as diagrams of familiar constellations. In short, the antiquity of maps is well established.

Not only are maps the oldest form of charting, but to this day they are the most widely understood. And the subject-matter they portray is of the greatest variety. Few, even of those who use maps regularly, have any idea of this diversity. Of the United States alone, there are on the market special maps showing the natural resources, the density of the population, the location and amount of the various crops, the chief centers of the various industries, the lines of communication and transportation. Some maps show political divisions, others the physical contours, others the mineral subsoil, and

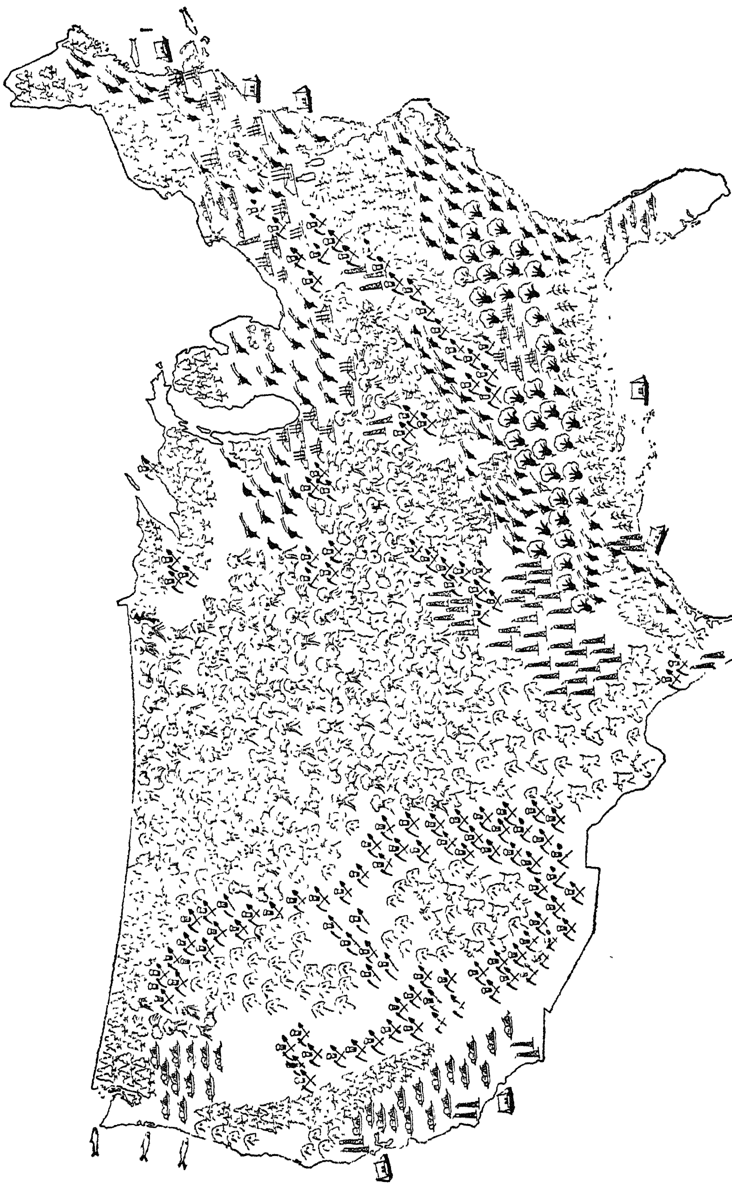


Fig. 1. Pictorial Map of the United States.

By Mr. C. Van de Wall.

still others the atmospheric conditions. Some show railroad distances between cities, others show automobile distances. It would be difficult to find any important phase of American life for which somewhere a map is not being published and marketed.

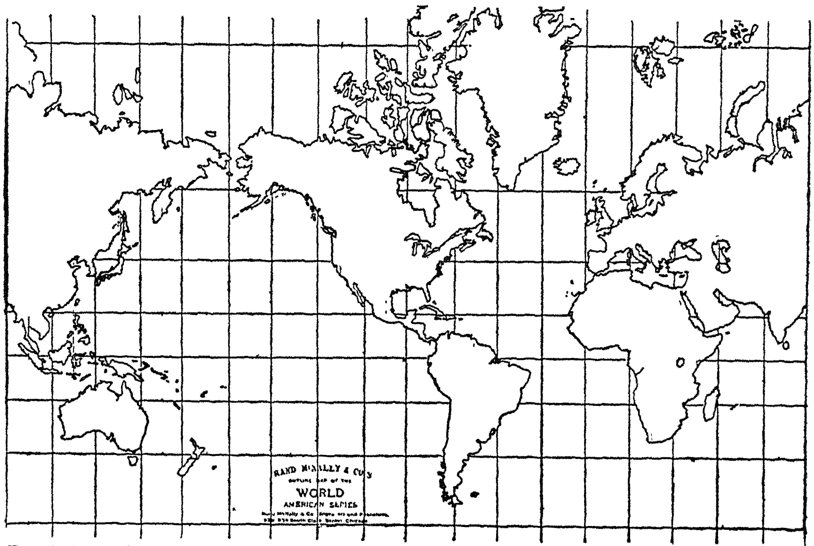
The student of maps will note that wherever large sections of the earth are shown, the map seems to suffer a distortion of outline, so that two maps of adjacent territories will not fit closely together and form a single large map. He will recognize that this is due to the fact that the earth is a sphere, while the map is printed upon a flat surface. We are indebted to one Christopher Columbus, who proved that the earth is round, for the necessity of this distortion. The result is that only upon globes can outlines be truly represented. All flat maps being more compressed, as it were, in their centers, and expanded at their edges, the outlines are consequently warped. So, too, it follows that maps of large areas, such as the United States, differ considerably in shape, according as the map is an imaginary picture of the country from a position above its southern, northern, or other parts.

In maps of the world, this distortion problem has become

From Bartholomew's Atlas.

Fig. 2. Heart-shaped Map of the World.

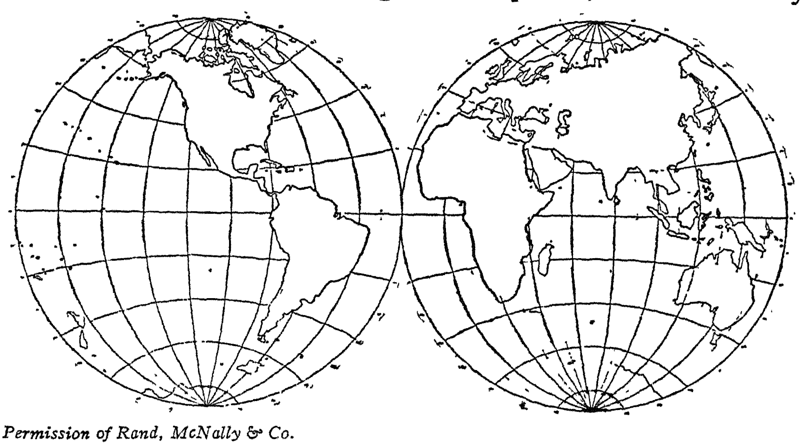
well-nigh insuperable, for world-maps must show us at once all of the earth's surface. Imagine seeing all sides of an apple at once! The most usual representation meets the difficulty



Permission of Rand, McNally & Co.

Fig. 3. Mercator's Projection of the World.

by magnifying the polar regions and spreading before us the sides of an imaginary cylinder. As the earth is not a cylinder and its poles are not as long as its equator, but are merely



Permission of Rand, McNally & Co.

Fig. 4. Hemispherical Projection of the World.

points on its surface, the amount of distortion can be seen to increase gradually from the equator and to become infinitely

great at the two poles. But by increasing the longitudinal or north-and-south dimensions equally with the intersecting latitudes, local outlines over small sections of the map are reason-

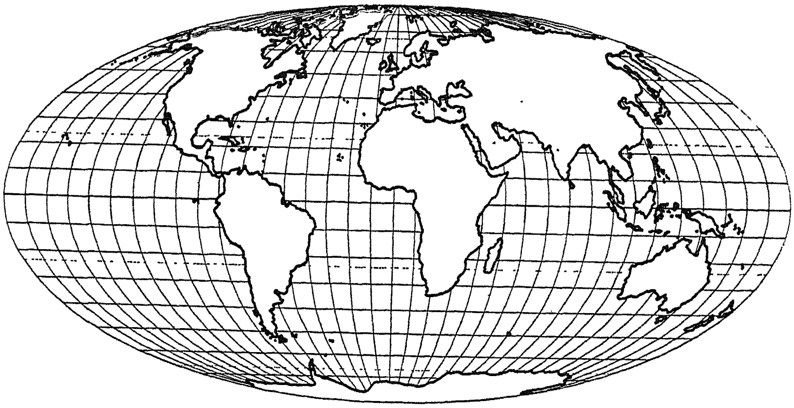
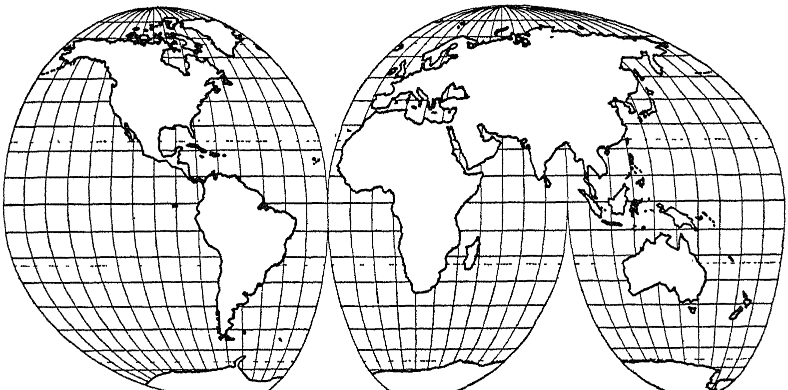


Fig. 5. Elliptical Projection of the World.

ably preserved. This device is called "Mercator's Projection."[1] Only when areas at unequal distances from the equator are compared, does this map become grossly deceptive. Who can forget his earliest impressions of Greenland being larger than Australia, or his amazement at the size of Canada and his wonder at the enormous reaches of Alaska, as gained from his world-map at the beginning of his school atlas?



Adapted from maps of the Hammond Map Company

Fig. 6. Homolographic Projection of the World.

[1] Invented by Gerardus Mercator, a Flemish mathematician and geographer (1512–1594), in 1550.

Such projections, of course, have no uniform scale of miles, for the inch that represents a thousand miles at the Desert of Sahara will represent but a few miles near the North Pole. A different form of distortion, a combined skewing and warping, takes place in the less common maps shaped either in two circles or in one flattened circle or ellipse. The best preservation of true outlines and areas, that is, a more uniform map-scale, is secured in a recent form of world-map called the "orangepeel projection," while for less broken outlines the "butterfly-map" and the homolographic projection[2] may be found useful. These are ingenious devices to keep recognizable shapes, each gaining its advantages only at the cost of simplicity and continuity.

The problem of distortion due to the earth's curvature tends of course to disappear as the areas chosen for presentation on the map become smaller and smaller. In large state maps it is still present, but the problem in county maps is rarely seen, so that several county maps can be fitted exactly together. Likewise a "scale of miles" holds true throughout the map when the area is small. Township and city plans are maps of still smaller surfaces and, indeed, the category of cartography[3] is not complete until we include floor-plans and diagrams of buildings and rooms, and the like. These are familiar in the form of architects' blue-prints and differ from maps proper only in that they can be quickly prepared by anyone, their subject-matter being of such limited space as to require no professional engineering surveys. However, they are in principle the same as maps, in that they are likewise representations of space in the plane of the earth's surface.

These plots, plans, and diagrams of small areas are so often of great value that we shall here explain in detail how they may be made. The first step, of course, is to secure the information to be charted. Assuming that you wish to draw a floor-plan, select some convenient point of reference, such as, perhaps, a certain corner of the elevator-shaft, and from this point of reference measure the distances to the various objects you wish to show on the plan. Measure these distances not

---

[2] The Encyclopedia Britannica, for example, lists some twenty-five different projections for maps of the world, of which the most distinctive have been here described.

[3] Cartography, according to the Century Dictionary, is the art or practice of drawing maps or charts (that is, marine maps).

directly to the objects, but along lines parallel to the sides of the room. Thus a certain motor stands, say, fifty feet east and ten feet north, of the corner of the elevator shaft. Take a piece of paper on which the objects to be shown have been listed in a column, and enter these figures beside each item, in

Plan of Small Statistical Office

	East from outer door	North from outer door
Chair of Statistician	8	6
" " First clerk	-2	5½
" " 2nd clerk	-5½	5½
" " Draftsman	-13	5½
" " Typist	-12	2
etc.		

Fig. 7. Data for a Floor-plan.

two columns. In the first column, headed East and West enter "plus 50" beside the motor ("plus" meaning "east" and

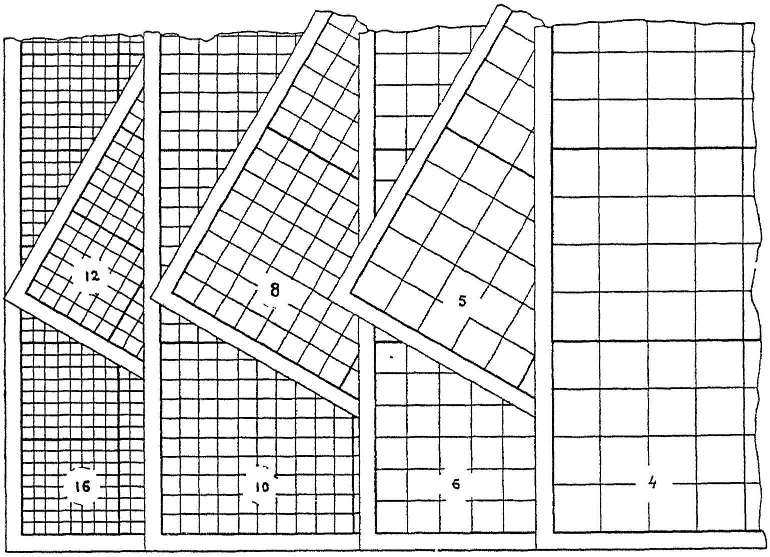


Fig. 8. Samples of Cross-ruled Paper.

The small numerals indicate the number of spaces per inch. Many other rulings are published.

"minus" meaning "west" of your point of reference, the elevator shaft). In the second column, headed North and

South enter "plus 10" ("plus" in this column meaning "north" and "minus" meaning "south" of your point of reference).

To draw a floor-plan or diagram, since no distortion problems[4] arise in such small areas, ordinary cross-ruled or "quadrille" paper may be used. Having prepared your data,[5] you will next decide upon a "scale" or ratio of reduction to use in the drawing, that is, what value or distance on the actual floor shall be represented by each space or distance between lines on the paper. It is important to pick a scale which is neither too large nor too small, so that the drawing will be the right size on the sheet. Suppose your paper is ruled in tenths of an inch

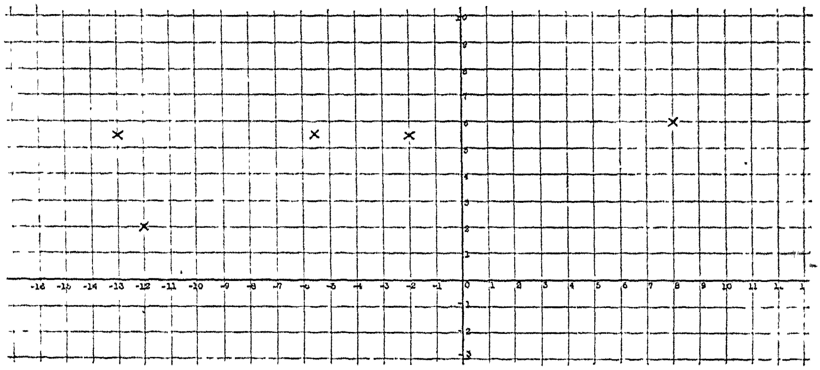


Fig. 9. An Unfinished Floor-plan.

with heavy rulings every inch, and you decide to let each small space represent one foot on the floor, and each inch ten feet. At some central spot on your paper where two heavy lines cross, mark the letter "O" to represent your point of origin. This point of origin on your paper corresponds to the point of reference on your floor. Along the heavy line through this "O" or zero-point, to the right mark the successive heavy cross-lines "10," "20," "30," and so on to represent distances east of the elevator, and to the left mark them successively " −10", " −20," " −30," and so on to represent westward direction. Along the vertical heavy line through the zero-point

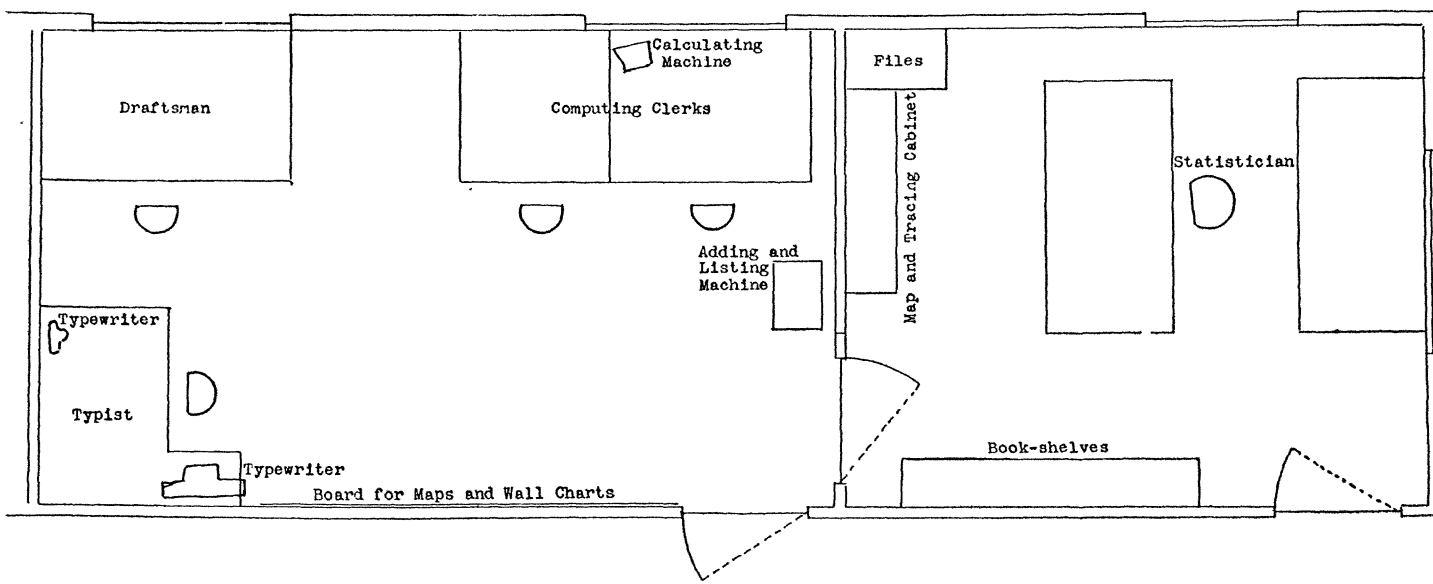
[4] The distortion in very large buildings may amount to several inches difference between horizontal distances of top and bottom floors, but does not appreciably affect the rectilinear outlines of floors.

[5] The word "data" is used throughout this book as a singular noun, unless it refers distinctly to more than one body of statistics. Such a usage is not sanctioned by the dictionaries, but is believed to be more in accordance with modern practice in the statistical work-rooms.

at right angles to the last, mark off the inches upward successively "10," "20" and so on to represent northward measurements on the floor, and " −10," " −20," and so on downward for southward measurements on the floor. After this, it is a simple matter to locate and draw on the paper each item in the spot corresponding to its true position on the floor. The scale-numbers "0," "10," etc., can be erased and the words "Ten feet to the inch" or a short calibrated line, substituted.

Here you have all the elements of chart-making. It only remains to observe the nomenclature. Our first step having been to secure two sets of measurements for each object or item, one for east-and-west distances and the other for north-and-south distances, we may call these measurements individually "values" and collectively "series." Where, as in this case, there are two values for each item, let us call one of them the "x" value and the other the "y" value in order to distinguish them easily. If a third measurement or series of values were present, it might of course be called the "z" value. In the present instance we have two, and, as will be seen below, the east-and-west series has been taken as the "x" series and the north-and-south series the "y" series. This is only a happen-so; we might equally well have reversed them, but as it is, we can now write for convenience the letter "x" over our first column of figures and the letter "y" over the second. So much for our data. It consists of two series of values.

Now on the chart, the two lines crossing at the zero-point or point of origin are called the "axes." The horizontal one is the "x-axis" and all values of the "x" series are measured along it, the positive ones to the right and the negative ones to the left of the origin. Parallel lines above and below this axis are called "abscissas" or "abscissae" and the axis itself is therefore sometimes called the "axis of abscissas." The vertical axis is called the "y-axis" or "axis of ordinates." Along it the values of the "y" series are measured, positively upward and negatively downward from the origin. The vertical lines parallel to it are called "ordinates." It will be noticed that all points on an abscissa have the same value of "y" and all points up or down an ordinate have of course the same values of "x". Taken together as a criss-cross pattern of lines, the abscissas (or horizontals) and the ordinates (or verticals) are called the "co-ordinates" of the chart.



FLOOR-PLAN FOR A SMALL STATISTICAL DEPARTMENT

Fig. 10. The Floor-plan, Finished.

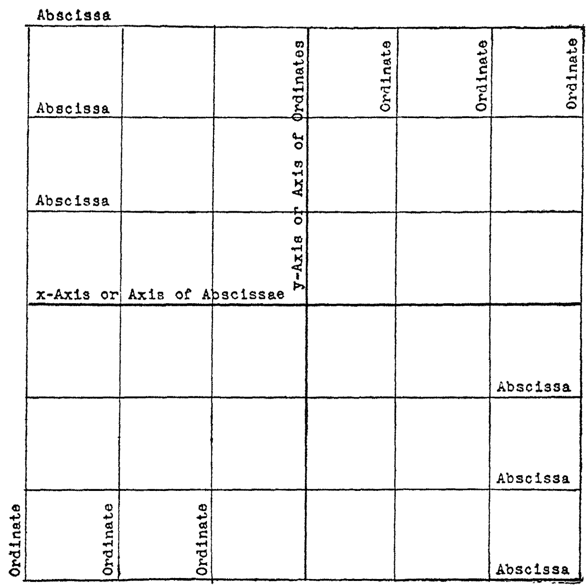


Fig. 11. The Nomenclature of Co-ordinates.

The student will observe that every point on this paper has two values, one along each axis, and that to identify or locate a point both its values must be given. He will observe that the axes cut the paper into four quarters (or quadrants), in the upper right-hand one of which (in our diagram the north-east quarter) both values of every point are positive, while in the lower left-hand quarter (south-west) both values are negative, and in the two other quarters one value is positive and the other negative. He will observe that, disregarding plus and minus signs, at each side of either axis, the values along the other axis always mirror themselves.

Many American cities are laid out in this checker-board style. In New York the north-and-south roads are called avenues and the east-and-west roads streets. In Washington the former are designated by numbers and the latter by letters. In both cases the house-numbers began at certain axial roads and read away in both directions. The conception is the same as that of the system of co-ordinates in the chart. Nor is it changed when we place the point of origin at a corner instead of in the center, that is, restrict the chart to one

quadrant, and thereby eliminate mirrored duplication of values and the need of plus and minus signs. This is commonly done in commercial maps, each map having a series of letters and numbers about its edges, the letters on two opposite sides and the numbers on the other two, each locating positions on one of the two axes. In the index or list of cities on the map, corresponding to our data, the proper combination of letters and numbers for the map is given to enable us easily to find any particular place. In short, the thoughtful reader will see that the fundamentals of charting are already familiar ideas, and will not allow a less familiar terminology of axes, abscissae and ordinates to confuse him.

CHAPTER II

CLASSIFICATION CHARTS

From the portrayal of space-relation between objects, we turn naturally to the portrayal of idea-relations, and to the relations of abstract ideas having no space-existence. Instead of location on an actual surface, we wish to show position in a more or less ideal scheme. We now deal, not with a geographical, but with a logical analysis. It is not possible to illustrate all the uses of charts in diagrammatic logic, but the classification chart is sufficiently suggestive.

In all chart-making, the material to be shown must be accurately compiled before it can be charted. For an understanding of the classification chart, we must delve somewhat into the mysteries of the various methods of classification and indexing. The art of classifying calls into play the power of visualizing a "whole" together with all its "parts." Even in the most exact science, it is not always easy to break up a whole into a complete set of the distinct, mutually exclusive parts which together exactly compose it.[1] A child can tell us that the United States is a single nation (whole) composed of forty-eight States and a District (parts), but almost everyone will find difficulty in deciding the number of territories, possessions, and spheres of influence which also compose it.

A second problem arises when each of the parts is in turn considered as a whole and its own parts analyzed.[2] Thus the State of New York is composed of 62 counties, that of Maryland consists of 23 counties and one city, and the counties are

[1] The division of a whole into many parts is sometimes called polychotomy; dichotomy and trichotomy are cases of division into two and three parts.

[2] The decimal classification is a case of repeated subdivision in which decimal figures (with or without the decimal point) are used as symbols or keys to the parts. The Dewey decimal system for book libraries is a familiar example of this method and the expansion thereof by the Brussels Institute Internationale de Bibliographie, founded by Senator Henri La Fontaine, is the greatest achievement in classification the world has ever known.

variously divided into townships, boroughs, incorporated places, and so on. Even a child knows that a dollar is theoretically divided into ten dimes, each dime into ten cents, and each cent into ten mills. But no two botanists agree in the classification of flowers, for example, into families (wholes), genera (parts), and species (sub-parts), not to mention the elaborate hierarchies of orders, classes and divisions, and the multitude of sub-species, sub-sub-species and hybrids.

The classification chart clearly presents, however, just so much of this marshalling and regimentation of ideas and objects as its author has clearly in mind. It is a method of presenting his scheme of things instantly and interestingly. Let us assume that he has settled his classification, has reduced it to writing, and tabulated it with indented margins or some other device to make it clear, and let us proceed to the technique of its charting.

The simplest form of chart showing a whole and its parts and sub-parts is the *box-chart.* It is composed of squares, rectangles, circles, or other "boxes" arranged in serried ranks down its page. Across the top, a single very large box carries the name of the total group (whole). In a row beneath it and

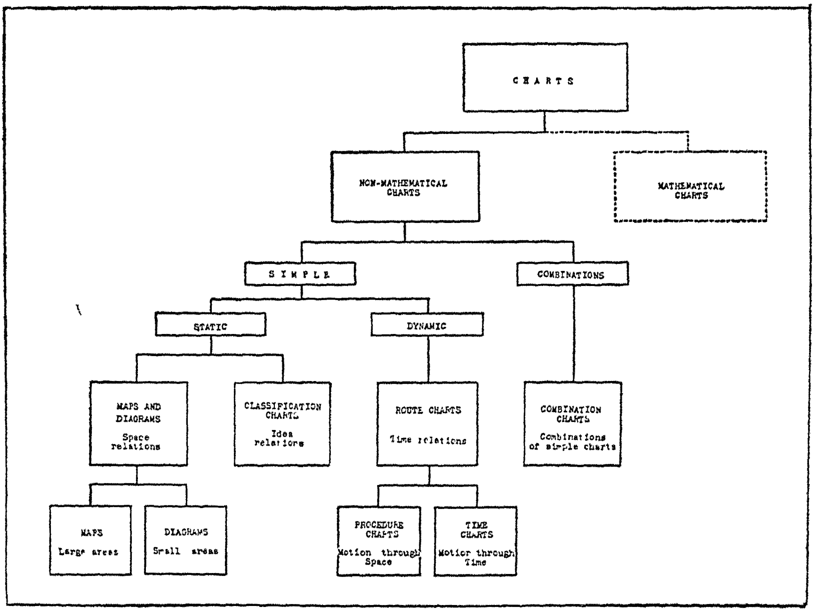


Fig. 12. A Simple Box-Chart.

tied to it by connecting lines are several smaller boxes, each bearing the name of one of the primary subdivisions (part or sub-total).[3] Beneath these again is a row of still smaller boxes, each similarly connected to one of the boxes in the row above and labelled with the name of one of the secondary subdivisions. The process may be continued indefinitely downward, to subdivisions of lower and lower rank.

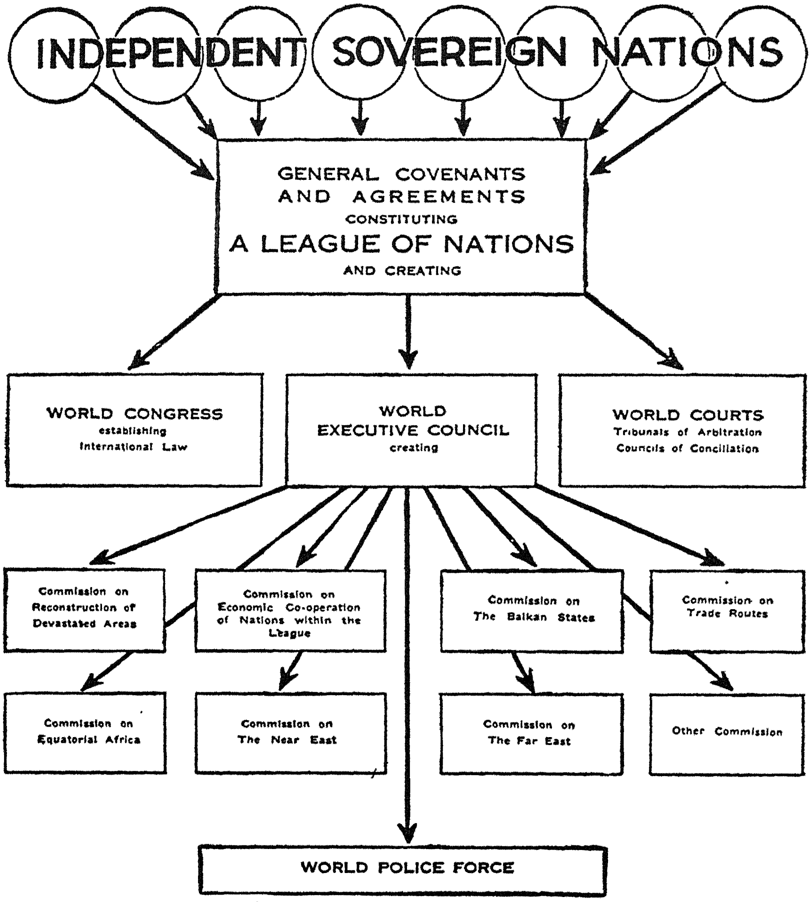
Sometimes there are so many subdivisions that they cannot all be shown side by side. In this case there are three courses open to us. The method most frequently employed happens to be the least desirable. It consists in dropping some of the minor boxes down to lower levels and connecting them vertically with the boxes above them. The method is unsatisfactory because it complicates the reading of the chart, changing the significance of a lower positioning on the page. When this method must be employed, it is well to distinguish the different ranks by various shapes, sizes or colors of boxes.

A second method is preferable. It consists in using very deep and narrow boxes for the minor subdivisions which must be crowded together. The labels will, of course, have to be written downward in these boxes; they can, however, be hung diagonally so as to make the reading easier. A third method is an outgrowth of the second. In it the entire chart is thrown over upon its side. The main or total box now appears at the left of the page instead of at the top: and the process of subdividing is carried out to the right, each rank in a different column.[4] This method is limited to cases of few ranks. Both the second and third methods are sound in principle, the significance of relative positioning being adhered to throughout.

The square or rectangular type of box is the best, being the easiest to draw and the clearest to read. Often it is perfectly feasible to omit the boxes entirely, taking care to keep the printing in box-formation. Where two or more distinct classes of objects are thrown together in a single chart, such as persons and departments, it is a happy thought to give one shape of box, such as a circle for persons, to one type of object, and a totally different shape of box, such as a square for departments

[3] The word "sub-total" is here used in its strict sense as an inferior or subordinate total, a part of the grand total which can itself be viewed as a whole and split into parts. It is not used in the customary accounting sense of a cumulative.

[4] If a mathematical chart showing the value of each of the final subdivisions is desired, the bar-charts described in a subsequent chapter may be used in conjunction with a classification-chart in this form.



Courtesy of Mr. Sidney Gulick

Fig. 13. Chart with Boxes of Various Shapes.

to the other type. Such differentiations should have a definite purpose, however, and must not be introduced merely to embellish the chart, as they then invariably complicate its reading.

Connecting lines may be either curved, straight, or rectilinear. The last are usually by far the best, especially where the boxes are rectangular. Straight lines, running directly between the boxes, give a radiating effect and are sometimes good when the boxes are circular. Curved lines generally fall into the class of pointless and undesirable embellishments, but are occasionally useful in complicated charts to connect boxes across other connecting lines. Ordinarily, when the lines cross, small semi-circles at the intersection on one line suffice.

In drawing the boxes and connections, full continuous lines will naturally be used, but it sometimes happens that certain parts of the chart are only remotely related to the main body of the chart, or perhaps belong to a different period of time. In this class, fall contemplated future additions to the existing scheme. Here the use of broken or dotted or even wavy lines is of value, not only for connection-lines, but also for box outlines. Another means of differentiation, discussed later, is the use of color or shading. It is somewhat more diverting to the

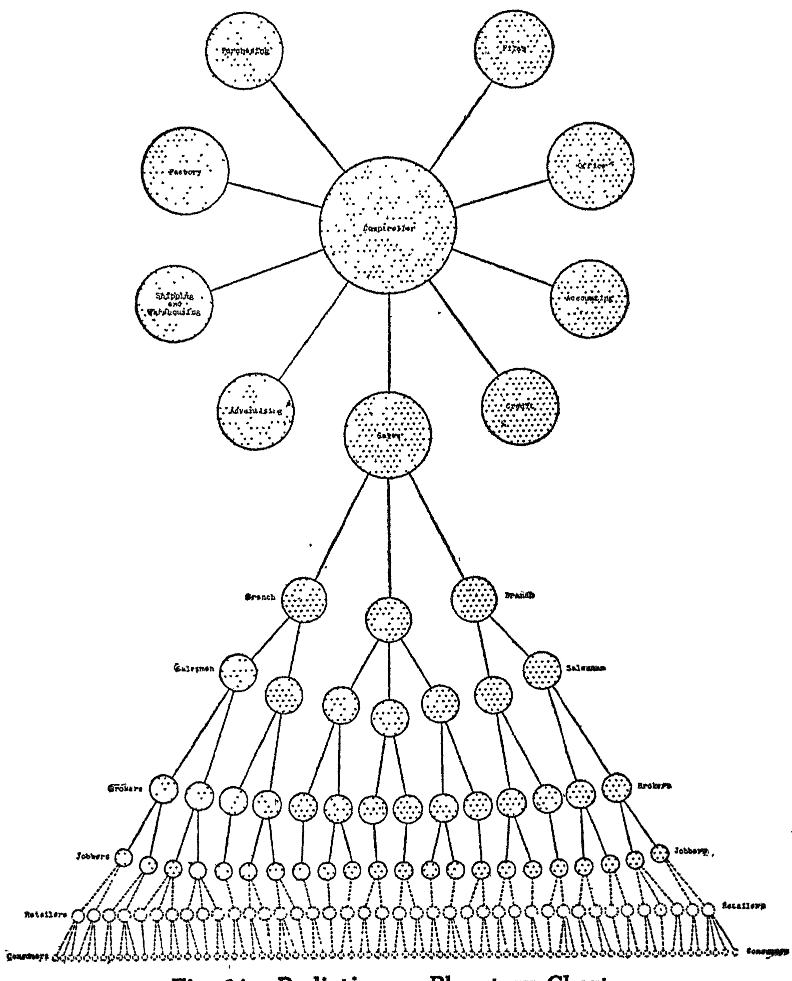
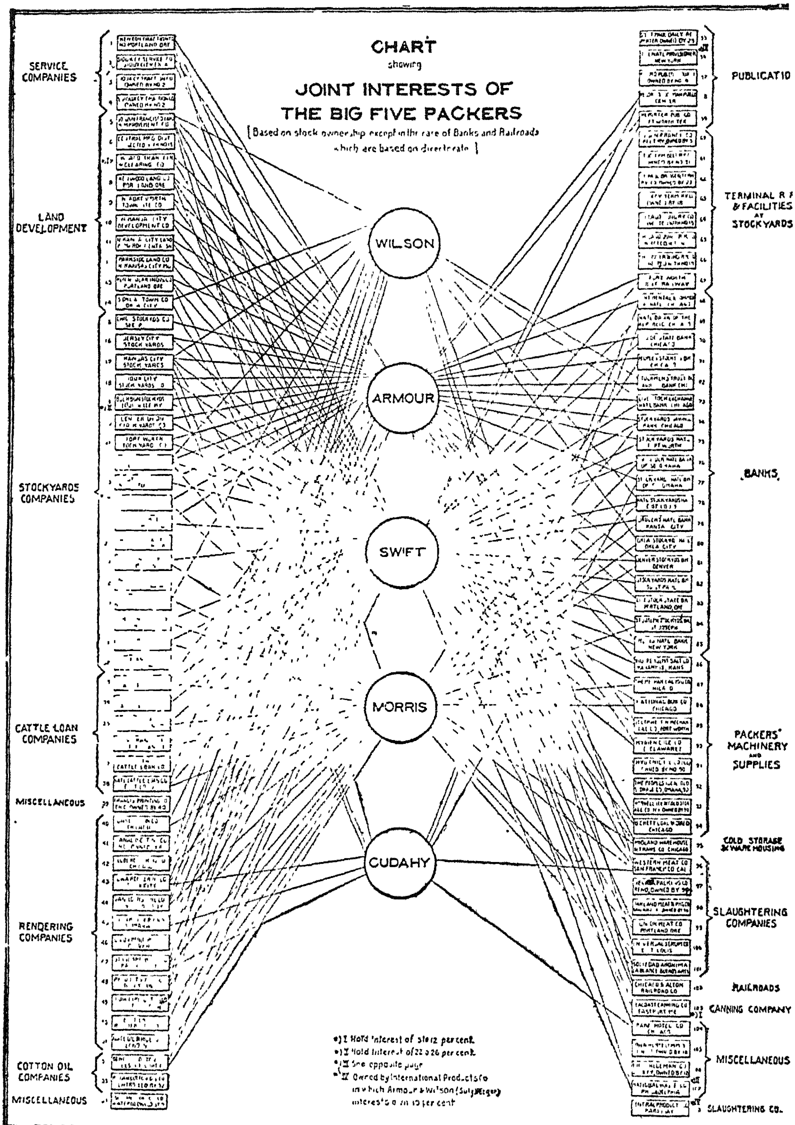


Fig. 14. Radiating or Planetary Chart.
Showing the Structure of a large Merchandizing Organization.

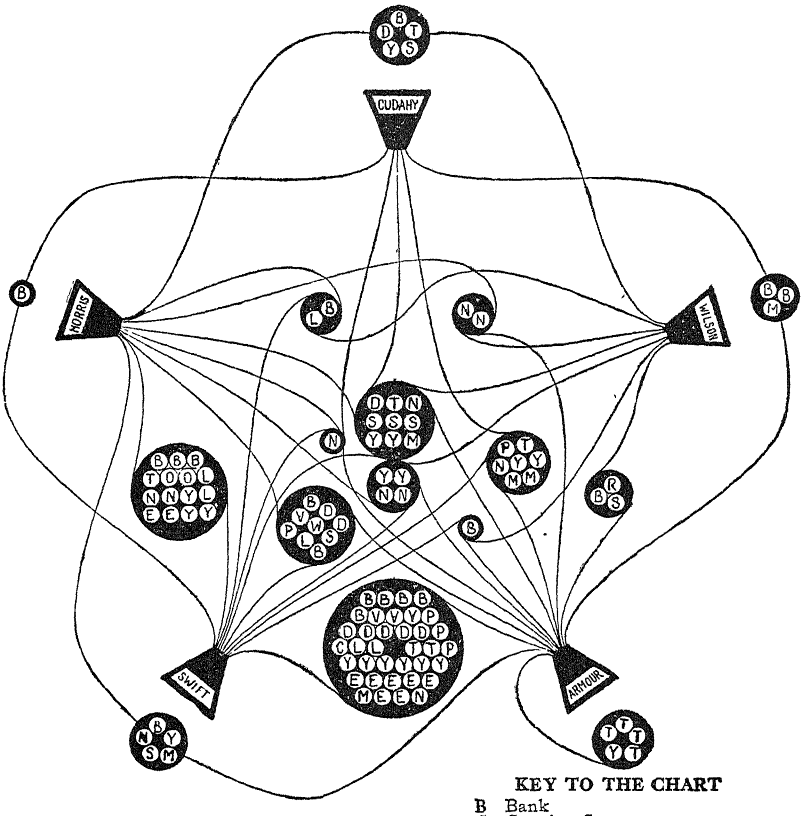


Reprinted from "Summary Report of the Federal Trade Commission on the meat-packing industry, July 3, 1918."

Fig. 15. An Example of Complicated Data.

reader, a dubious advantage, however, and it cannot well be reproduced.

The variations of the box-chart which are occasionally seen



KEY TO THE CHART

B Bank
C Canning Company
D Land Development Company
E Packing Equipment Company
L Cattle loan company
N Miscellaneous
Rendering company
O Cotton oil company
P Publishing house
R Railroad
S Slaughtering company
T Terminal railroads and facilities at stockyards
V Public service companies
W Coldstorage and warehousing companies
Y Stockyards companies

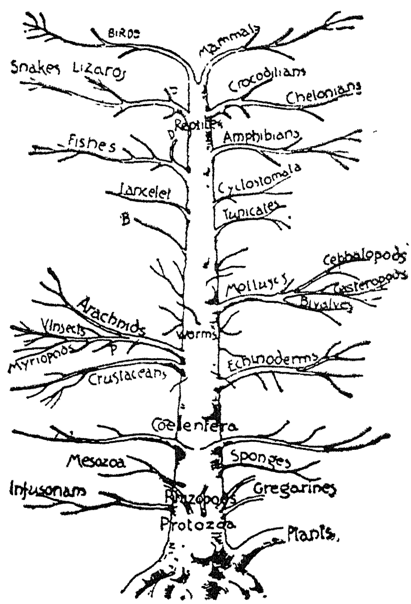
JOINT INTERESTS
OF THE
"BIG FIVE" PACKERS

(*Source: "Summary Report of the Federal Trade Commission on the meat-packing industry July 3, 1918."*)

Fig. 16. Five Interlocking Classification Charts.

are usually inspired solely by a desire for an artistic appearance and raise in turn some doubts about the material they present, that is, doubts as to its accuracy or the spirit in which the chart was compiled. The "tree-chart" is a sample of this, in which the trunk of the tree represents the total group, the branches the primary subdivisions, and the leaves, twigs, or fruit the

minor subdivisions. Another is the "planetary chart," in which a central sun is labelled the whole group, its planets the primary subdivisions, and their satellites, either encircling them or outside of them, the minor subdivisions. Such variations are justified only when something in the nature of the object shown suggests a particular fitness in the figure. The variation always makes for a certain amount of difficulty in



From Thompson's "Outline of Science," published by G. P. Putnam's Sons.

Fig 17. Tree Chart.

Showing the Evolution of Animal Life.

reading the chart, it takes a great deal more time to prepare, and, if well done, is likely to draw more attention to its own arrangement than to the subject-matter it is intended to convey.[5]

[5]The student will detect in classification-charts certain elements in common with the diagrams described in the previous chapter. Axes are no less present because they are not drawn and calibrated; for in one direction the positioning signifies lower subdivision, while in the cross-wise direction it signifies equal and independent importance. No scale is used because the variables are not numerical measures, but only idealogical relations. While the analogy is not important, it is interesting to keep in mind.

Chapter III

ROUTE-CHARTS

To the executive type of mind few charts make such instant appeal as those describing movement—that flow of goods through a sequence of operations which is the keystone of industry. Economics itself is but the study of the successive forms of "wealth" through the processes of production and distribution. Static relations, either physical, as shown in maps, or logical, as in classification charts, may engross the academic interest; indeed a correct conception of them is essential. But when through them is woven the added element of time and motion, the result is lifted out of the field of cut-and-dried research and given the values of life itself. And he who weaves such a pattern performs, no matter in how small a way, a creative engineering function. The picture of such a process we call the route-chart. This chart throws powerful light on the weaknesses and advantages of a process, either existing or contemplated, and gives to the reader, perhaps even to the author, a grasp of the subject which no amount of text can equal. It is a photograph capturing that highest of human achievements, the mental visualization of action.

As is often the case in chart-making, preparing the data for this chart is no small part of the work. The data consist of the accurate record of the steps, changes or events which take place. This record may be compiled in the form of notes or text. In simple cases the successive steps may be listed or tabulated, using indented margins where the process branches or splits into different channels. Such data are very similar in form to the data for classification charts, already described. But where the process is complicated with detours, by-products, cross-connections and detailed assemblings, no list or tabulation will remain clear and the data must take the form of a careful statement in notes, possibly in conjunction with card-indices, a cross-reference system and rough working sketches.

It is in fact not a bad practice in extreme cases to use a large bulletin-board or wall, and, having the information written on scraps of paper, to arrange and re-arrange these scraps of paper with thumb-tacks thereon, until the final order is settled.

Sometimes apparently complicated data turns out to be a series of the combinations and permutations of simple elements. Every step or event is then merely a combination of two or three or more items of a descriptive nature. When these descriptive component items are broken apart and listed individually, it will usually be found that they are few in number and can be grouped according to their nature into different series, in such a way that one item from each series is present in each event. Commonly, three of these component series are sufficient to identify all the events. A production process,

MAKING THE CURVE CHART

	(Subject)	(Operation)	(Operator)
1	Data, Sources	Securing of	Chief
2	Computing	Instructions for	"
3		Execution of	Clerk
4		Checking of	"
5	Chart, Data	Inspection of	Chief
6	Field	Choice of	"
7	Data	Entering of	Typist
8		Checking of	Clerk
9	Scale	Choice of	Draftsman
10	Curve	Plotting of	"
11		Checking of	Clerk
12	Scale	Entering of	Typist
13	Chart	Inspection of	Chief
14	Title	Choice of	"
15		Entering of	Typist
16	Chart	O. K.	Chief

Fig. 18. A Tabulation of Simple Route-chart Data.

for example, may be made up of operator, object, and operation ("who," "what" and "how"). Time ("when") may also be actually recorded. A distribution process may be made up of combinations of place, person and proportion ("where," "by whom" and "how much"). The number and nature of these component series will vary widely in different processes,

but the above are fair samples. Needless to say, where the data can be analysed in this way, it will simplify the work of compilation and assure the completeness of the data to list the events, together with their component details, in parallel columns, a column for each series or type of detail.[1]

A still more condensed type of work-sheet can be prepared for complicated data, in which only two types or series of descriptive items are present. This consists of a diagram in which each series is listed fully and once for all, along an axis

MAKING THE CURVE CHART

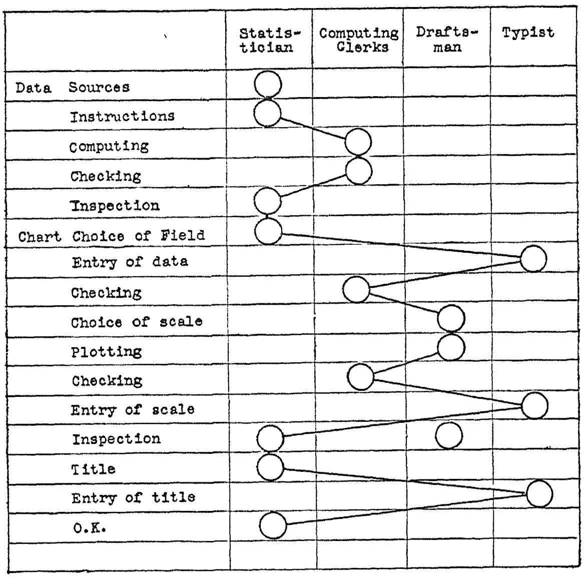


Fig. 19. A Condensed Work-sheet.

or edge of the paper. Thus on paper with columnar rulings, list the descriptive items of one type or series (e.g. materials) down the left-hand edge of the paper, and those of the other type or series (e.g. departments) across the tops of the columns. Then along the line of each first-series item mark with a cross or circle the columns of those second-series items with which it is combined to make an event or step (e.g. operation). It makes little difference which series is put along either edge.

[1] The student will be reminded of the keys and data-sheets used for maps and diagrams, in which the two sets of values, or measurements along the axes, were listed in column form, and will compare the descriptive detail series to the numerical value series.

If one series is longer than the other, it should be at the side, but if both are of the same length, have the more important series—the series by which events are to be grouped—listed at the side edge of the paper. If this be reversed and the more important series placed across the top, the crosses or circles would of course be entered up and down the columns instead of across them. The sequence of events can be shown by small numerals in the circles, or by connecting lines with arrowheads, or best of all, by both. These connecting lines

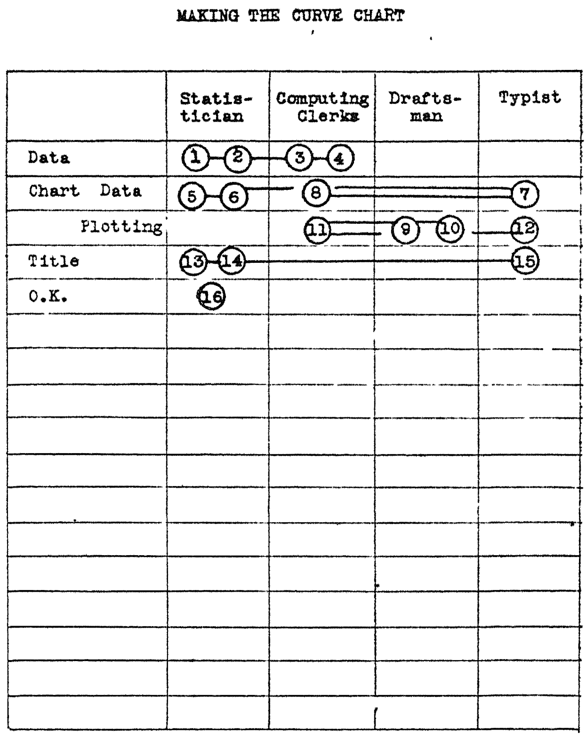


Fig. 20. A Very Condensed Work-sheet.

showing sequence would run horizontally if the side-edge items are more important, vertically if the top ones are more important. In practise either arrangement is satisfactory and it is usual to leave the longer series listed down the page, because more items can be written down a page than across it, even when the column-headings are entered on edge.

The phenomenon of motion involves two inseparable elements, space and time, in either of which the motion may be

measured. It is ordinarily simpler to prepare first an analysis of movement through space. This will include changes of location, condition or operations. It is, in fact, simply a measuring of events by arbitrary differences in their nature, instead of a measurement by differences in point of time. When it has been completed, it may also be desirable to measure the movements chronologically and to co-ordinate them upon the chart so as to show graphically the motion through time as well as space. Assuming that the data for the chart have been decided upon, we shall proceed to its graphic presentation, beginning with the simpler route-chart, showing only change of place or condition. For convenience, this may be called the "procedure-chart."

The simplest "procedure-chart" is a straight line or row of "boxes" with the steps or events inscribed and with arrows

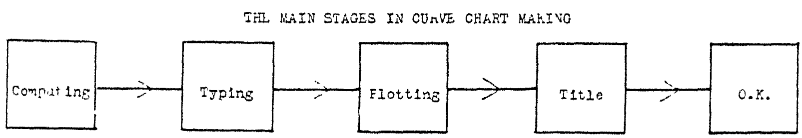


Fig. 21. The Simplest Procedure-chart.

along the connection-lines indicating the direction of movement. It is important that the arrangement of the steps be in a uniform direction across the paper. They can be arranged horizontally from left to right, or vertically from top to bottom (or even, in special cases from bottom to top). The same considerations will determine this direction as were noticed in the arrangement of the classification-chart; namely, the lettering of the boxes generally gives them greater breadth than height, and if there are but a few boxes, they can be placed side by side. However, if there are many, they can be packed closer one above the other.

The procedure-chart is in many respects similar to the classification-chart, its main differences being that it need not branch or split up at each new step, and that the connecting lines between boxes indicate a path-way or line of motion. For more complicated data in which the processes branch out and split up, the similarity between the two charts will be very great. The use of different styles or shapes of boxes now becomes more advantageous, as the various steps may be totally dissimilar, and by adopting certain shapes for each

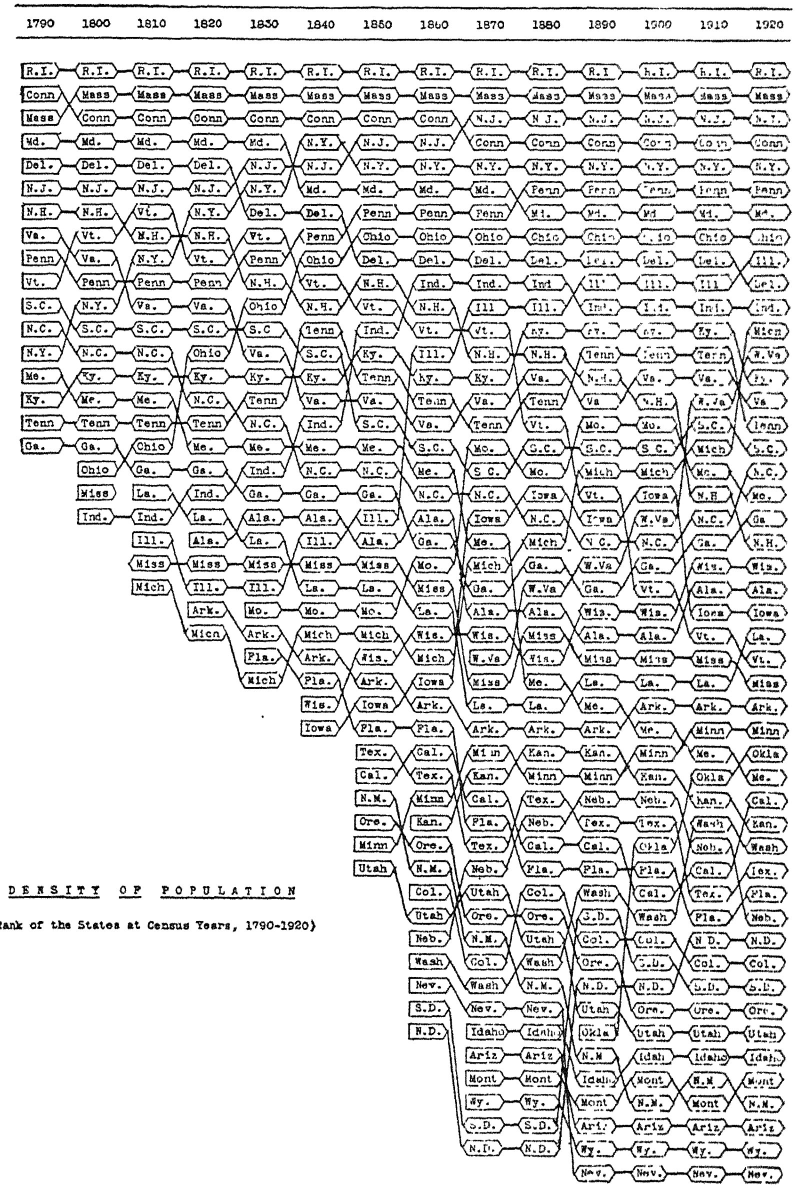


Fig. 22. A Simple Procedure-chart with Many Items.

type of step (e.g. operation) or for each type of descriptive detail (e.g. departments, persons, objects, functions, etc.), these distinctions between steps or events are clearly brought

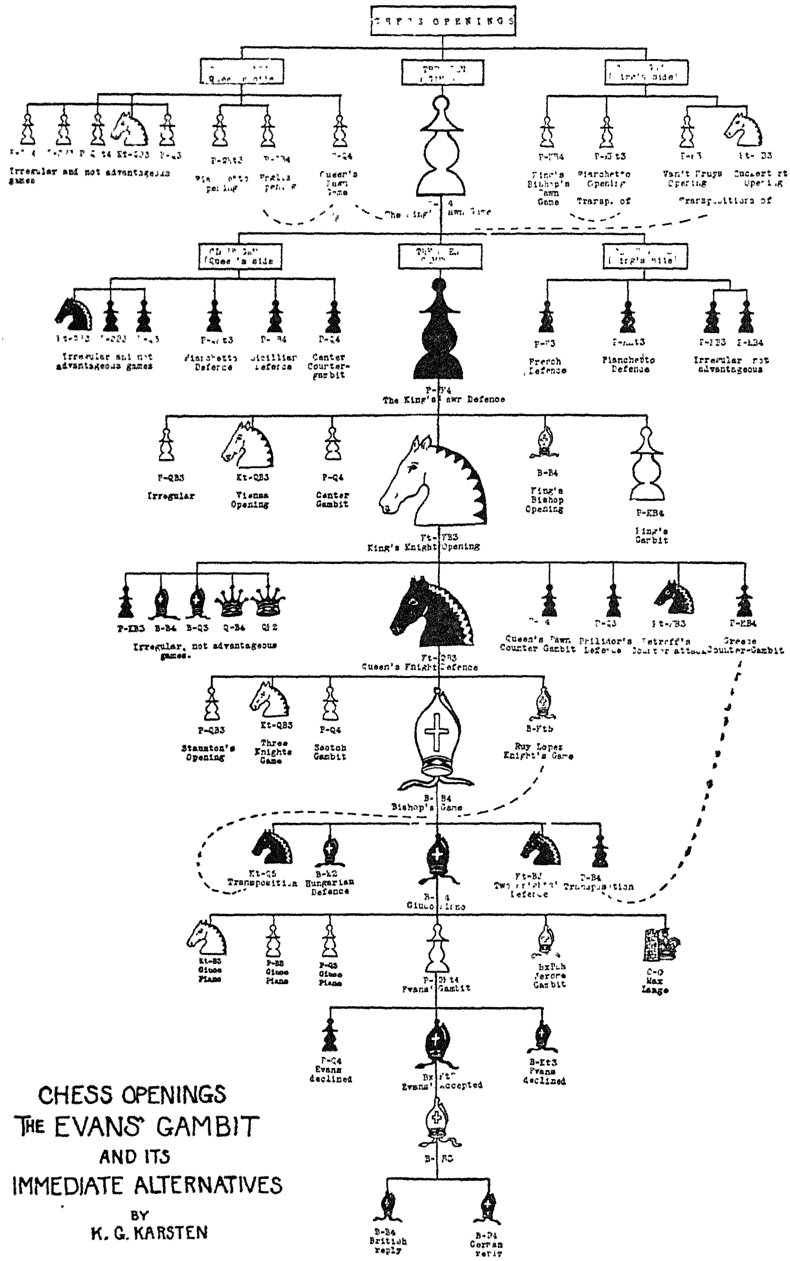
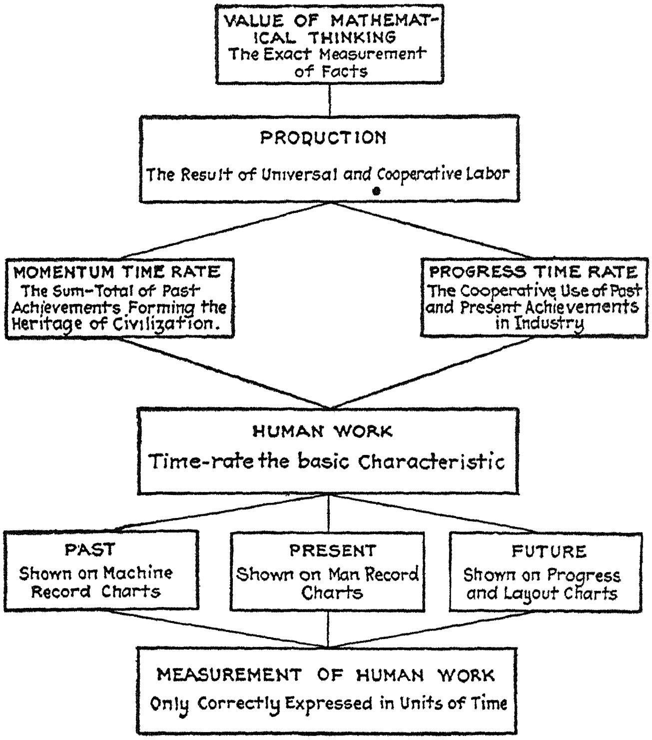


Fig. 23. A Pictorial Procedure-chart.
In which the objects or materials of the various steps are realistically pictured.

out on the chart. Indeed, it is sometimes possible to add greatly to the value of the chart by picturing the various



Showing the subjects treated in an article by Walter N. Polakov, in Engineering Management, by permission.

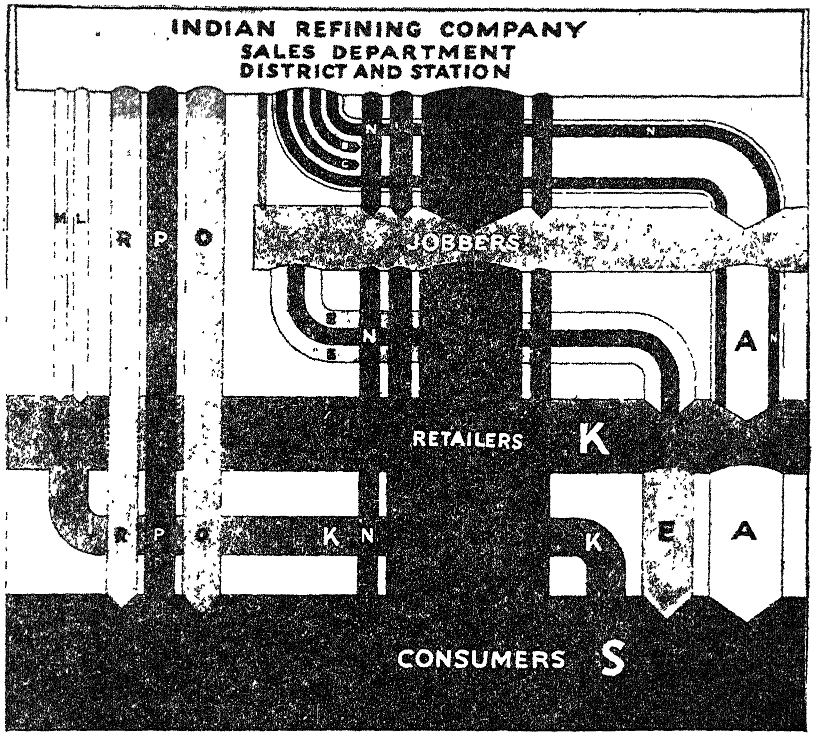
Fig. 24. A Graphic Outline of Thought.

stages realistically. At other times the chart is so simple that the boxes can be omitted entirely.

A uniform direction of movement across the chart is important, because it automatically suggests to the reader the sequence of events. It would better be described as a uniform drift; motion at right angles to this drift, necessary at branchings of the process, being immaterial. There are occasions when, on account of the data, it is necessary to draw a line backward, as is the case when seconds or by-products return to an earlier stage for re-treatment, but these are legitimate representations, suggesting actual backward steps of the process. At other times it is necessary to choose between backward directions of lines and repetitions of boxes; as a rule the latter is the lesser of the two evils, but if the former is decided on, the backward motion should be strongly indicated by arrows, and the connecting-lines should leave boxes and

enter boxes at the points they would naturally leave and enter if the boxes were in proper sequence.

Embellishments, artistic and otherwise, are often met with. Impartial study will usually show that nothing has been gained by them, and that the message of the chart would have been



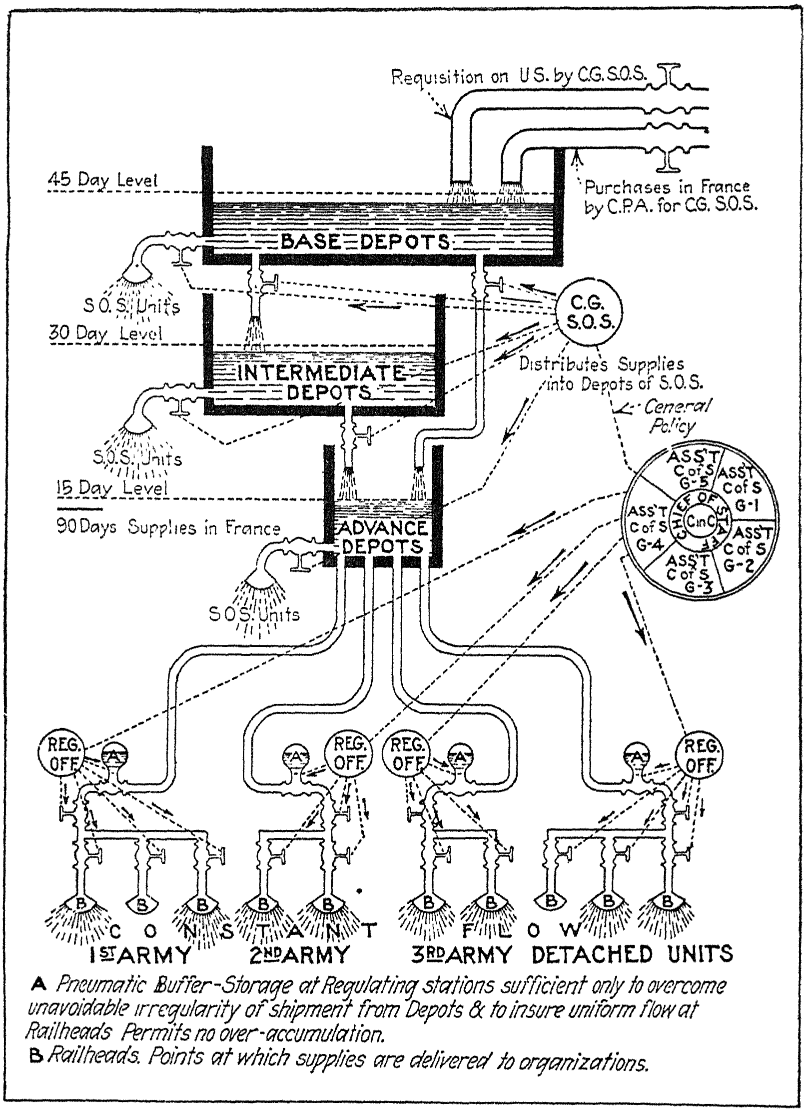
Permission of Mr Richard Webster

Fig. 25. A Popular Presentation.

The original of this chart, prepared in colors, is provided with a key explaining the various channels through which influence is brought to bear upon consumers by the sales department.

more strongly conveyed without them. The occasional exceptions to this rule are special cases of data in which the subject-matter itself suggests the modified form as particularly appropriate. In the "boiler-chart" the source of supply is shown as a large tank or boiler, and the goods are shown as flowing through pipe-lines to smaller tanks, cylinders, engines and outlets, each representing a particular type of comparable stage in the process. Here the reader's imagination is fired by the implied simile of a familiar mechanical process, and if

the simile be a good one, he is likely to examine it closely and so visualize the process clearly. Sometimes a row of tanks or

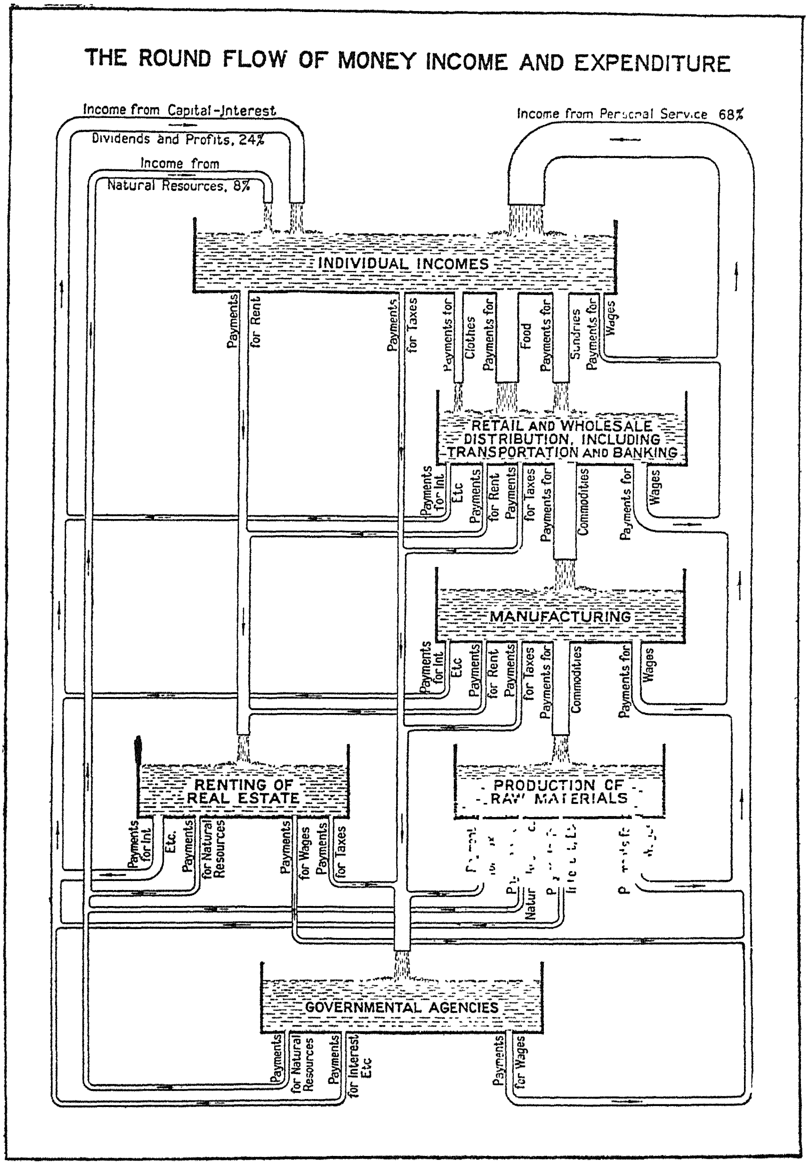


Permission of Mr. Malcolm C. Rorty.

Fig. 26. An Excellent Pictorial Route-Chart.

This shows the flow of supplies in the American Expeditionary Force.

vats will represent successive cost-burdens well, the overflow (profit or balance) from each flowing into and feeding the

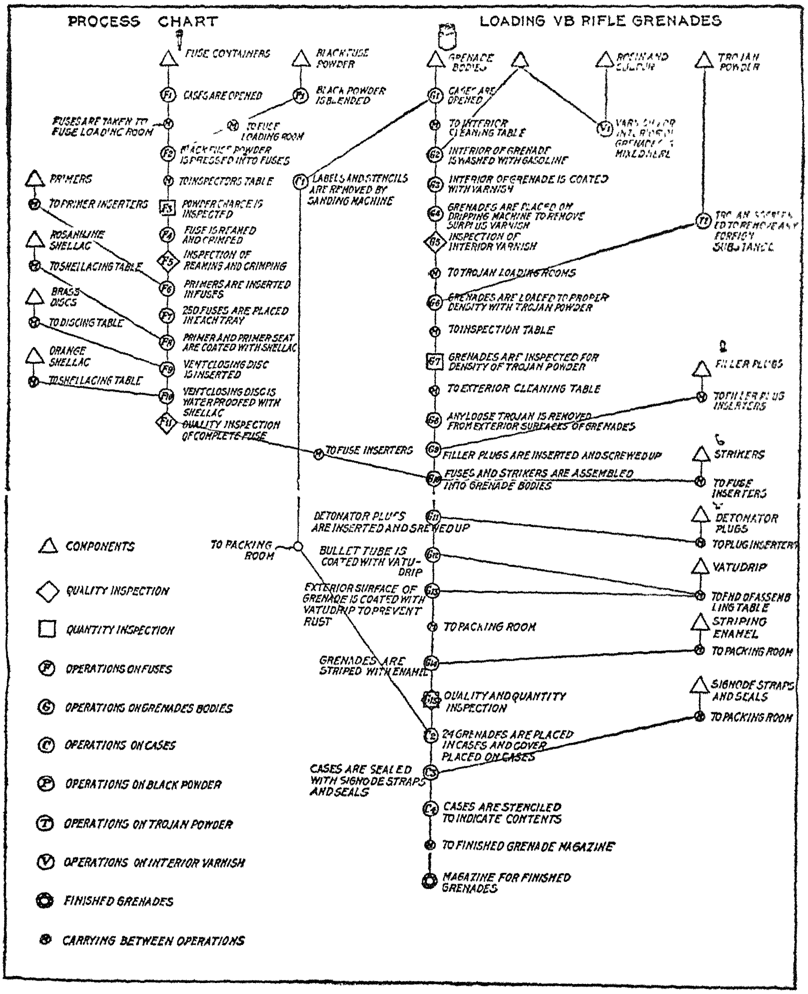


Permission of Mr. Malcolm C. Rorty.

Fig. 27. The Analogy of Vats, Tanks or Reservoirs.

next. There is no limit to the possible variety of such representations, but in the main the chart-maker will find the widest play of his imagination called for in the making of legitimate

procedure-charts, and will do well to avoid the excessive and wasteful effort needed for the artistry of more sensational products.



From "Process Charts and Their Place in Management," by Frank B. Gilbreth and L. M. Gilbreth "Mechanical Engineering," Jan. 1922.

Fig. 28. A Simplified Gilbreth Process-Chart.

Showing operations by conventional symbols and materials by pictorial drawings.

The second type of route-chart differs from the first or procedure-chart in that time is an important element in the data and a feature of the chart. It may be called the "time-chart." It is easily made by arranging a time-scale along the

axis or direction of movement of the procedure-chart, and adjusting the various boxes or entries of events so that their positions coincide with the ordinates or abscissae of their particular points of time. The scale of time should be marked along the edge of the chart (in large charts along both edges)

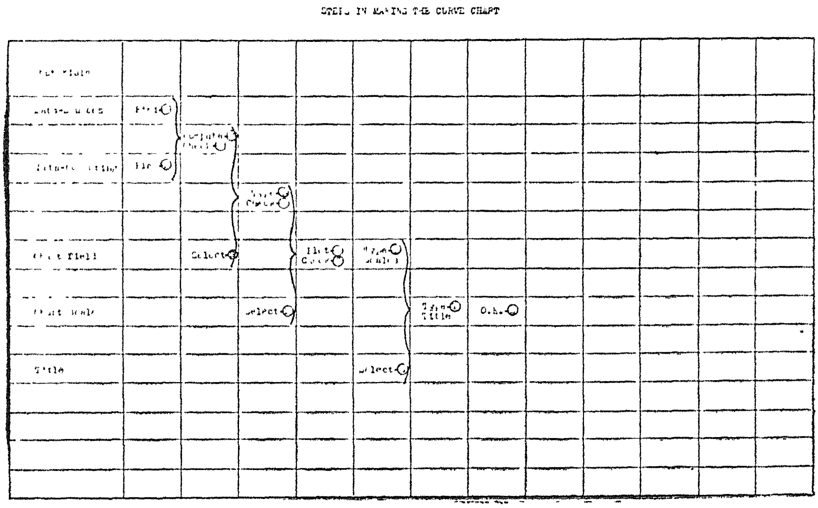


Fig. 29. A Simple Form.

In practice this chart would bear dates for the days or weeks or other time intervals at the top of each column across the page.

and straight lines in a faint color should be ruled across the chart from the main divisions on this scale. A glance across the chart will then show, by means of these faint lines, how many time-intervals elapse between steps or events and a study of the scale will give a more exact estimate when desired. In time-charts, needless to say, the uniform direction of movement or drift is essential.

The time-chart may be reduced to a form similar to that of the work-sheet already described for procedure-charts. With the time ruled off on one axis of the diagram, the items to be followed by the chart are listed on the perpendicular axis and the action of each item indicated by crosses, checks, or solid shadings along the line of this item under or opposite to the right moment of time. By various kinds or colors of shadings, or by words alongside the shadings, the nature of the event happening to the item can be indicated. In principle, it is better to place time on the horizontal axis, so that a standard time-scale can be used and long charts folded in

Curves -- 1/20 to 2/5/17

CURVE-CHART PROGRESS RECORD

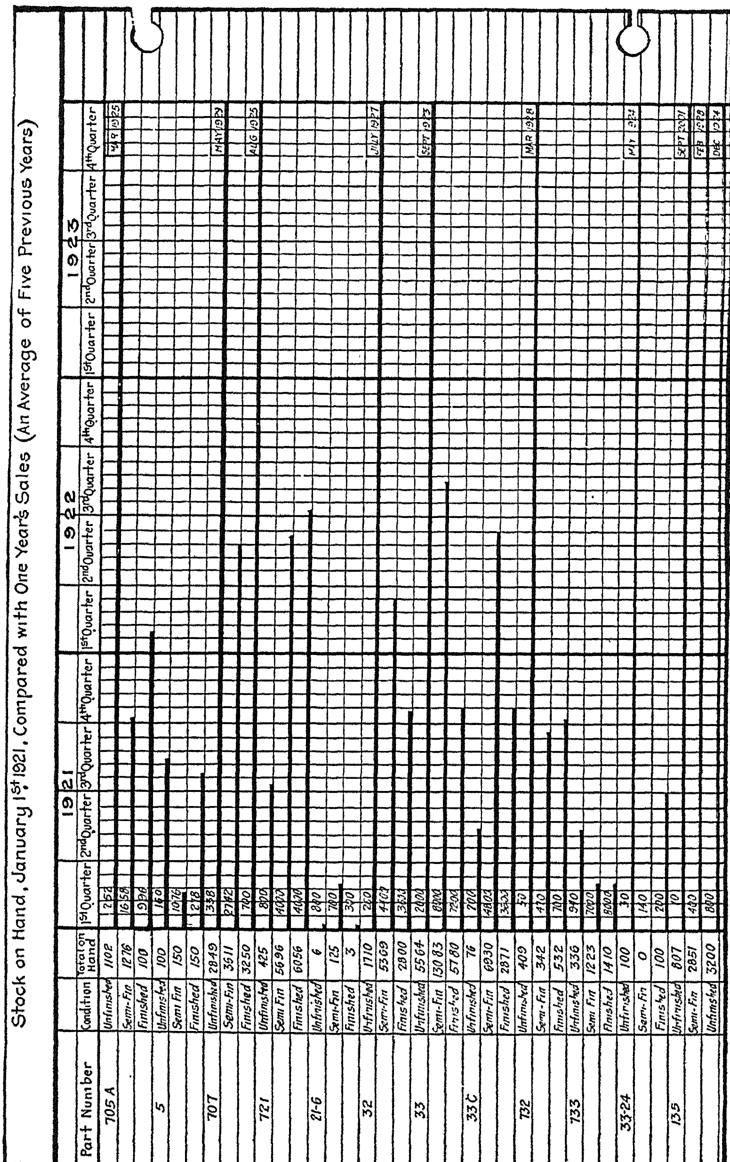
CHART	NUMBER OF REPORT	NUMBER OF CHART	SOURCE	COMPUTING PLAN	COMPUTING EXECUTION	COMPUTING CHECK	O.K.	CHART FIELD	DATA TYPING	DATA CHECK	CHART SCALE	CURVE PLOTTING	CURVE CHECK	TYPING SCALE	O.K.	TITLE CHOICE	TYPING TITLE	O.K.
Inventories	V-4	1	1/20	1/20	1/21	1/22	1/22	1/22	1/23	1/23	1/24	1/24	1/24	1/24	1/25	1/25	1/25	1/26
		2	1/20	"	"	"	"	"	1/22	1/22	1/23	1/23	1/23	1/23	1/25	"	"	"
		3	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"
		4	1/30	1/20	2/1	2/1	2/1	2/1	2/2	2/2	2/2	2/2	2/2	2/2	2/2	1/20	2/2	2/2
		5	"	"	"	"	"	"	"	"	"	"	2/3	2/3	2/3	"	2/3	2/3
		6	1/22	1/20	1/23	1/23	1/23	1/23	1/25	1/25	"	"	"	"	"	"	"	"
		7	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"
Shipments	SV-1	1	1/21	1/21	1/21	--	--	--	--	--	--	--	1/21	1/22	---	--	--	1/22
Warehousing	SM-7	5	1/25	2/10	2/10	2/10	2/10	2/10	2/10	2/10								
General	P-6	17	2/1	2/1	2/1	2/1	2/1	3/1	2/1	2/1	2/1	2/1	2/1	2/1	2/1	2/1	2/1	2/1
District 9 Sales	SM-3	14	2/2	2/2	2/10													
		15-8	2/4	2/4														

Fig. 30. A Time Record, But Not a Time-Chart.
This is inserted for comparison with the previous illustration, as the dates form the body of this table instead of the column headings, and the operations form the column headings instead of the body.

sideways to reduce them to the same size records or files. If very finely ruled paper be used, a large amount of detail can be crowded into this sheet and if a number of similar processes are to be compared, a standard arrangement of the items on it will make quick comparison easy. This simplified form of time-chart has little to recommend it from a graphic point of view, but it will be found extremely convenient and sometimes indispensable as a record, and is always useful as a work-sheet in preparing a more graphic time-chart.

An example of this simplified time-chart is the Gantt chart method. All charts in the Gantt system employ uniform vertical rulings, marking off "time" on the horizontal axis. The particular markings are always adjusted to the individual business, so that the spaces between vertical rulings may indicate hours, days, weeks, or months as is desired. And the columns between these vertical lines, after their adjustment to the periodicity of the particular business, resemble columnar accounting sheets. At the left-hand edge of the paper, in a very wide preliminary column, the machines, departments,

materials, or other form of equipment are listed. And along the lines of each of these items, under the proper "time" (that



From Wallace Clark's "The Gantt Chart," by permission of the Ronald Press.

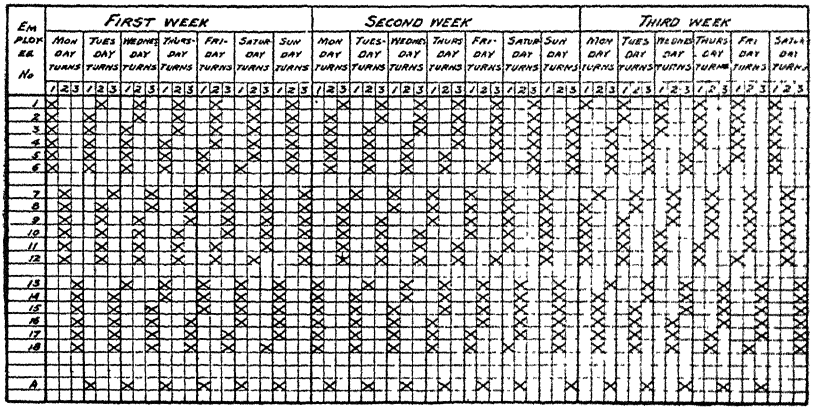
Fig. 31. A Gantt Chart.

Here the items are sufficiently indicated in the stubs and are graphically recorded (as to time) by bars drawn out to the proper point of time.

is, in the column of the proper time unit) the process or operation is noted by a line commencing at the time of the beginning

of the operation and ending at the time of its finishing. This line takes the place of the crosses or circles above described and has the advantage of showing by its length the length of time required for the operation. The chart is used not only for laying out work in advance in the planning department and in the individual departments, but is also used to record actual performance when the records of the same are being kept. Wide lines are used in the place of narrow ones to record actual performance.

CHART FOR 8 HOUR TURNS WITH 24 HOURS OFF AFTER 6 DAYS. IN CONTINUOUS OPERATION INDUSTRIES



Permission of the Bureau of Labor Statistics.

Fig. 32. A Simple Time-Chart.

Another notable example of this type of chart is the Gilbreth "process-chart," used by Mr. Gilbreth to analyse methods of work in his well-known micro-motion studies. On paper ruled in tenths of inches, each fine line of which represents one-thousandth of a minute in time, the length of time to perform each operation or element of an operation is recorded by a heavy line, beginning at the commencement of the operation and showing by its length the number of thousandths of a minute required to complete the operation or element of an operation. Mr. Gilbreth ordinarily measures time vertically down the page, a detail in which the present writer believes him to be ill-advised. Across the page his vertical rulings mark off the different parts of the human body whose motions he is studying. There are about a hundred of these parts listed at the top of the page, each over one narrow column or space between vertical lines. Obviously the lines indicating

operations, extend vertically down the page under the parts of the body active in the operation. Different colors of the lines indicate different elements of the operation and different widths of the lines indicate degrees of activity engaged in the operation.

Where the time runs in natural recurrent cycles, such as days, weeks, months, or years, and items re-appear at identical points in each cycle, a circular form of time-chart is often

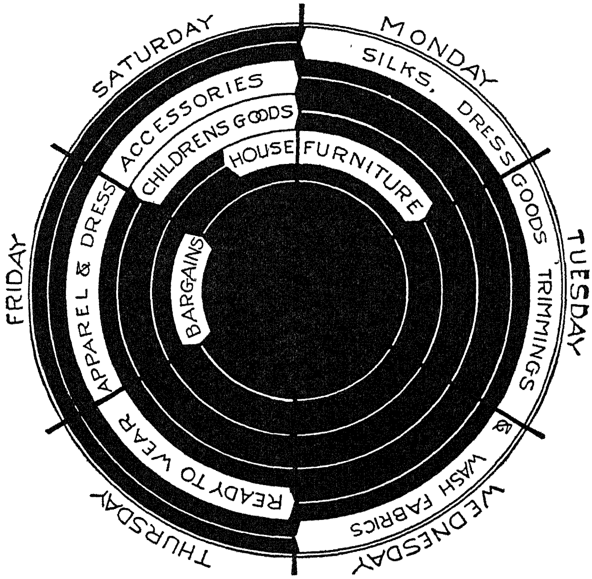


Fig. 33. A Weekly Clock-chart.
The weekly cycle of sales in a department store.

desirable. It has the advantage of being endless without actually showing more than a single cycle and naturally suggests to the reader the recurrent nature of the process. In such a chart the time-scale runs around the edge of the outer circle and the time-interval lines, equivalent to ordinates, appear as radii from the center of the chart. The items or events are inserted in boxes in their proper positions along concentric circles, those near the center having of course less room on the chart. Care should be taken to place the larger items, or events, requiring more descriptive labelling, toward the outside, if possible, in order that they may not be too crowded. Such circular time-charts are called "clock-charts" but they are not necessarily marked off like a clock; in one complete revolution they will show twenty-four hours for the day, or

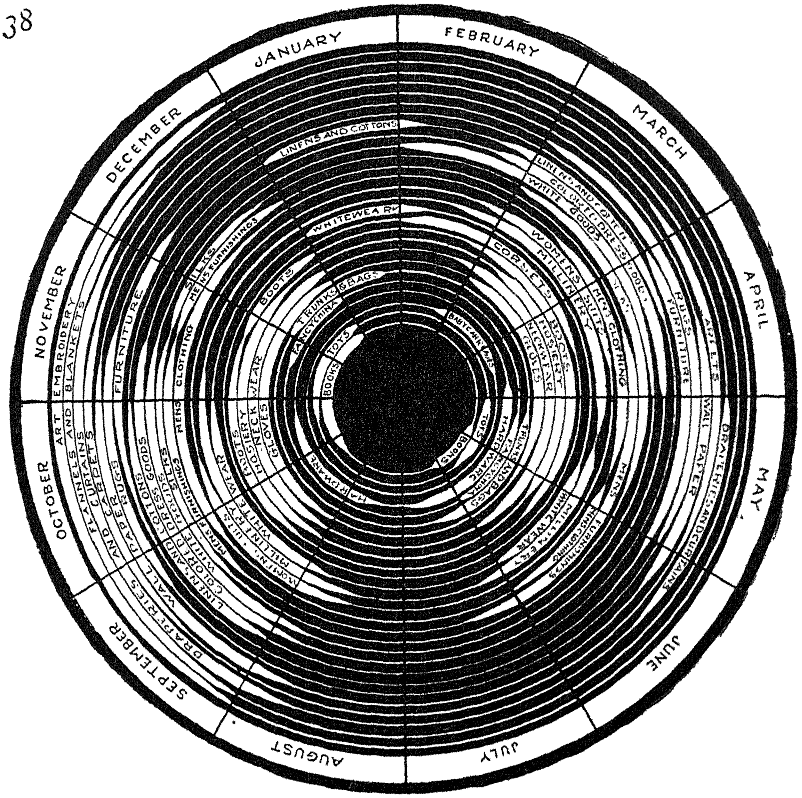


Fig. 34. An Annual Clock-chart.
The annual cycle of sales in a department store.

seven days for the week, thirty-one days for the month and twelve months for the year, according to the time-cycle chosen.

Route-charts, like classification-charts, offer great freedom to the ingenuity of the author. No set rules can be laid down for their construction, though the general principles above outlined will be found always safe and helpful. If the author of the chart desires to modify it, he will break no iron-bound canons, though he will probably have to do a great deal of experimental work before he has a satisfactory product. The one really final criterion by which his product will be judged will be, as in all charts, how clearly, forcibly, and truly does his chart tell his story. If he can pass this test better with a novel form of chart than with the typical and sound forms which have been described, he will have really invented a new statistical instrument and his product will be a contribution to the science, but the man with limited time will be well advised to follow and remain within the fundamental principles here outlined.

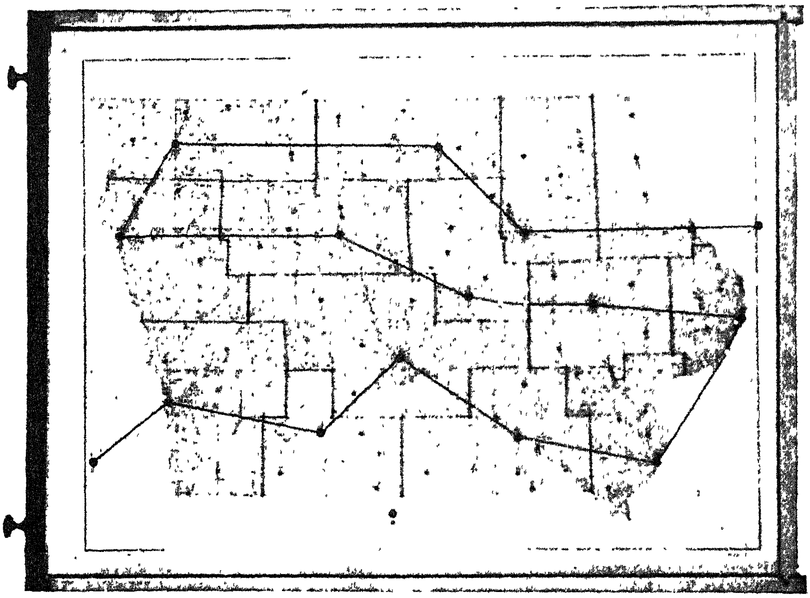
Chapter IV

COMPOSITE CHARTS

More fascinating than any one of the fundamental chart-types already described are the results developed by combinations of two or more of these types simultaneously. The simple types are three in number, adapted to showing between objects a space-relation (maps and diagrams), a topical relation (classification-charts) or a relation in motion (route-charts). Any two of these relations may be shown simultaneously by combining the principles of their chart-forms. It is only necessary to construct first one type of chart and then with this as a basis or ground-work, superimpose upon it the construction of another type, in such a way that while each retains its own significance, the two harmonize in details.

Very often the two will be so closely interwoven that they seem to be inseparable and indistinguishable, but the student will always find that under close analysis they readily break down into two or more separate and distinct charts belonging to the essential types which have been described. He will also find that this process of breaking down a composite chart invariably clarifies his understanding of its subject-matter, and vice versa, that the more obviously the component charts are distinguished in the composite product, the more clearly its subject-matter will be understood by its readers. If the chart is composite, it is important that the maker should recognize its nature, and it is important that the chart itself should show on its face that it is composite.

If you will take a map of the country through which you have travelled and with a heavy black pencil draw a line along the routes you have passed over, with circles or boxes about the names of places where you have stopped, you will have a simple form of a superimposition. Had you marked upon tracing paper over the map, instead of marking directly upon the map, you would be able to lift your second chart bodily

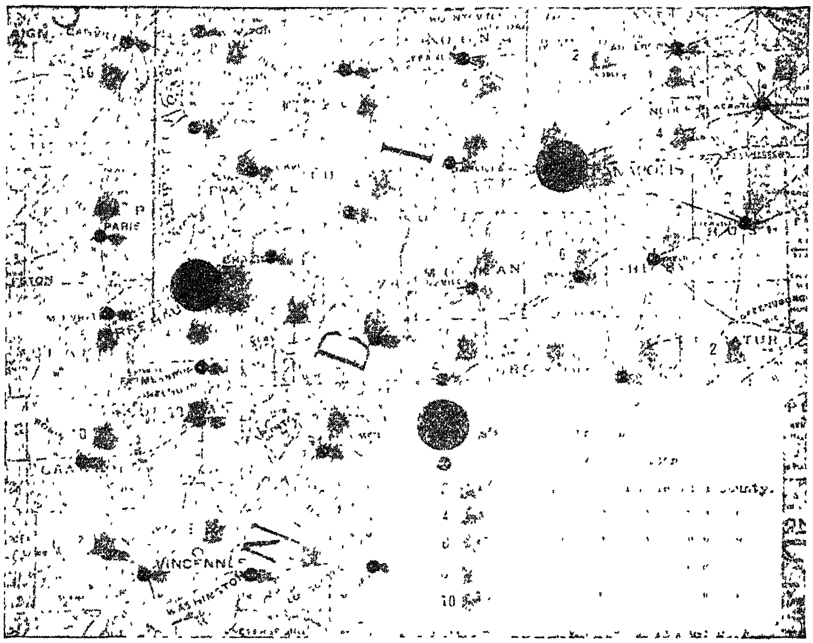


Courtesy of Rand McNally & Co.

Fig. 35. Route Map.

off of your base chart, and would see that you really superimposed a route-chart, describing motion, upon a map or chart of space-relations. The result is a composite chart illustrating motion through space.

Maps and diagrams are often used as a basis for charts showing motion. In fact there has recently been developed an elaborate technique of what are known as "pin-maps." These are particularly in vogue among sales-managers, who have to route a number of salesmen about the country and wish them to cover the most ground in the least time and with the lowest possible travelling expense. Maps for this purpose are mounted, and the markings upon them are made in the form of conspicuous colored tacks and other devices driven into or fastened onto the surface of the maps. In this form, the same map may serve for many temporary superimpositions and the latter can be readily altered at will, without the labor of complete re-drawing, merely by removing or shifting the adhesive markings. The labor-saving value of pin-maps is so great for all kinds of continual routing work that they are marketed in excellent form by various commercial firms, including nearly all map-making companies.



Courtesy of Rand McNally & Co.

Fig. 36. Pin Map.

The mounting of maps or diagrams to accommodate pins or map-tacks should be closely examined. Ordinarily the maps are mounted directly on wood, either to be framed and hung on the wall, or they are already fitted into flat drawers of special cabinets holding a large number of such drawers in horizontal positions. The wood-mounting is, however, a poor investment. Pins cannot easily be forced far enough into the wood to be secure, nor easily removed if driven deep, and sooner or later, as the wood shrinks under the punctured paper, individual pins will drop out, and cannot be replaced without complete rechecking of all data, the map meanwhile becoming inaccurate and unreliable.

The ideal mount for a map, and in the long run the cheapest, is a construction of cork and corrugated paper-board. The map should be mounted directly upon a piece of cork linoleum, with a non-wrinkling adhesive called rubber cement rather than with paste or glue. The cork should then be backed up with two or three layers of corrugated wrapping board or paper, laid in alternate directions to prevent bending. In a

mount of this sort, the pins can be easily pushed into the map to their heads; the cork grips them and prevents their falling out, and the corrugated paper keeps their points from sticking out underneath. Maps can be mounted in this way at home or in the office, and in some instances can be procured directly from the manufacturers of the maps.

Map tacks and other marking devices to attach to the mounted map can be obtained in great variety, suitable for showing a number of distinct markings and meanings at the same time. The tacks are small steel pins with large round or flat heads of cloth, celluloid, or, best of all, glass, conspicuously marked, in different colors and sizes. When they are inserted in the map to indicate, for example, towns on a salesman's route, the various colors can be used to indicate different salesmen, the various sizes can show the length of the salesman's visit, and the various markings on the tacks can tell the extent of the company's business there. Similarly, colored string can be stretched between tacks to show the sequence in which they are visited, while small celluloid rings of the same color can be slipped over the tacks to show present location or progress along the route. Ring-pins into which cards can be fastened, are made in various shapes to hold cards at different angles to the map, flat-headed pins on which labels can be pasted, rough-ground celluloided pins on which pencil-markings can be made or erased, and a wide variety of other appliances are furnished for ingenious uses with mounted maps.

It is, however, no longer necessary to have mountings and attached devices to make temporary and easily altered markings upon maps. Such arrangements take up too much space and require special filing or housing equipment if used extensively. Instead, it may be desired to use flat maps in book or sheet form, which can be more easily carried about. For this purpose a celluloid-coated map is made, the surface of which is protected by a thin adhesive layer of pliable transparent celluloid.[1] On this surface pen, crayon, or ink marks can be made without difficulty and removed without injury to the map, while gummed paper stars, dots, and other signals in various colors can be attached and removed likewise.

[1] The celluloided map is really little more than a map which has been surfaced with a thin layer of liquid shellac or varnish. The liquid can be secured from dealers in artists materials and can easily be applied to any chart or map with a fine blow-spray; it forms a protection against soiling, as the surface can always be cleaned without removing the marks under the coating.

The benefits of mounted maps suggest that floor-plans, and other diagrams could be likewise profitably mounted and used for pins and strings, but as a rule this is not yet a general practice. The pathway of goods, papers, or functions about a plant is generally marked upon photostats or blue-prints or other reproductions of one original floor-map, in various colored inks. The lines upon such diagrams should have frequent arrow-heads to indicate the direction of movement, as this direction is no longer shown by position on page as in the simple route-chart. Different colors or kinds of lines can be used to differentiate the pathways of various articles, but when a large number of such articles are to be individually followed, it is better to use a number of copies of the original base-map, one for each article or group of articles, so that the lines will not be too confusing.

An elaboration of this method has been described to the writer by Dr. C. W. Gerstenberg, who tells of "a factory where a bird's-eye view of each machine has been drawn to scale and fastened with brass paper-fasteners to a piece of cardboard cut to scale to represent the amount of floor-space needed for each machine. The color of the card-board indicates the nature of the machine; thus planers are on red card-board, drills on blue, and so on. The mounted machine-diagrams are then placed on a floor-plan of the factory and fastened into proper position (they can be shifted if the machines are shifted), and the routing of work is shown by ribbons slipped under the machines and stretched from machine to machine in the order of work. The flexibility of this ribbon idea commends itself."

It may be added that if several types of work were routed over this floor-plan, it might be advisable to use different colored ribbons to distinguish them. Furthermore, some idea of the volume of traffic or work along each route might be given by using ribbons of various widths, narrow ones for small or occasional work and wide ones for heavy traffic. The student will notice that, in accordance with the principles set forth below, the colors of machine-mounts should be in pale tints and those of work-routing ribbons in brilliant tints, to emphasize the route-chart over the floor-plan.

When several floors are to be shown on the same map, that is, when the diagram must show in one plane a number of surfaces which are really not side by side, but one above the

other, the co-ordinate or cross-ruled chart-paper can be discarded and a special form of paper, with "isometric "rulings, used in its place. This paper projects without any perspective,

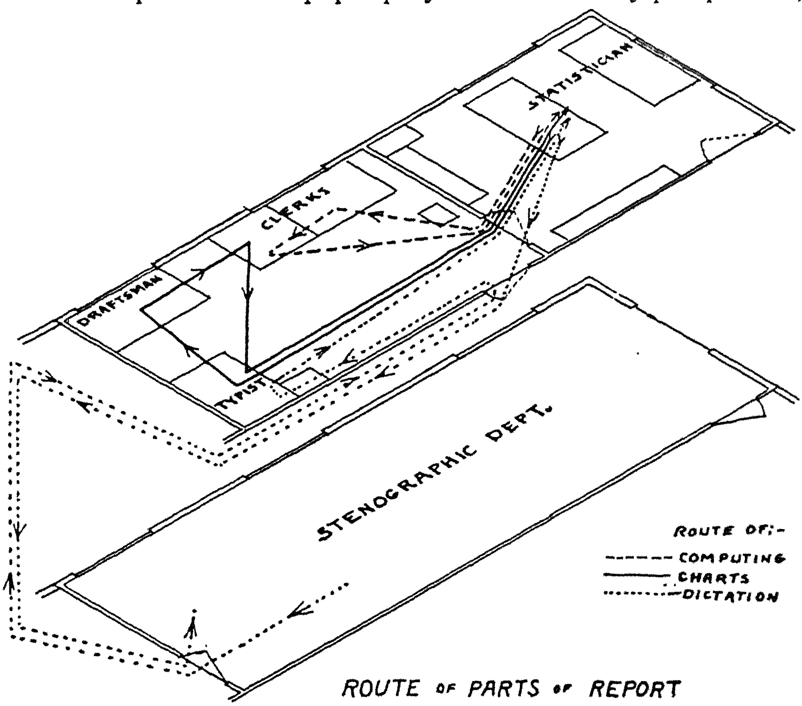


Fig. 37. An Isometric Drawing.

three dimensions in space, and following the general plan of charting on co-ordinate paper, can easily be plotted to show the various floors suspended one above the other. A full description of the principles of the paper will be given later, but the paper is noted here for its peculiar value in this type of composite chart.

Such route-diagrams often furnish the most forceful way of presenting the weaknesses of a given arrangement of a factory. Traffic congestion is apparent through the number of crossing and confusing lines. A contemplated change which will result in a more orderly march of goods about the factory will appear upon such a chart with all its benefits made clear. A knowledge of the chart-form will therefore be useful whenever a revision or improvement of the lay-out of the plant is in view, with the object of shortening transportation distances or of installing "straight-line" processes.

Many other composite charts may be made besides those showing motion through space. A classification-chart may be

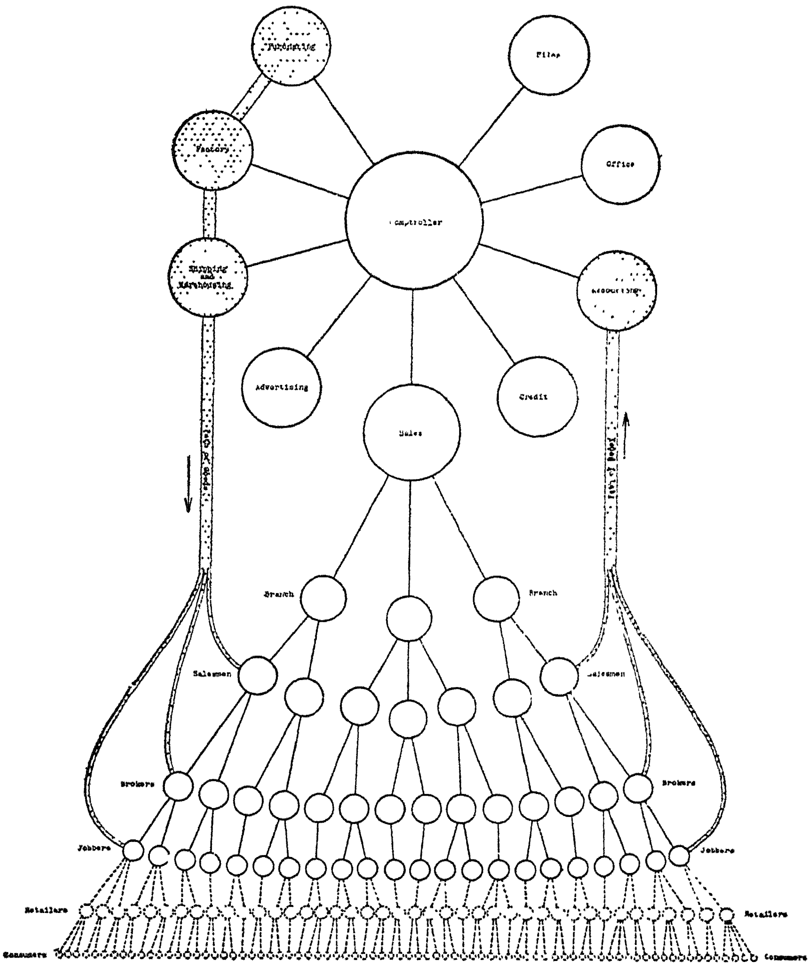


Fig. 38. Routing on a Classification-chart.
Showing by its shadings, the departments through which the circulation of goods and money takes place.

superimposed upon a map. A route-chart may be superimposed on a classification-chart. Two classification-charts may be combined to show in a single chart both sets of logical relations. Route-charts themselves may be combined to show routings at different times or may be marked with distinctly classifying features. There is no limit to the variety of ways

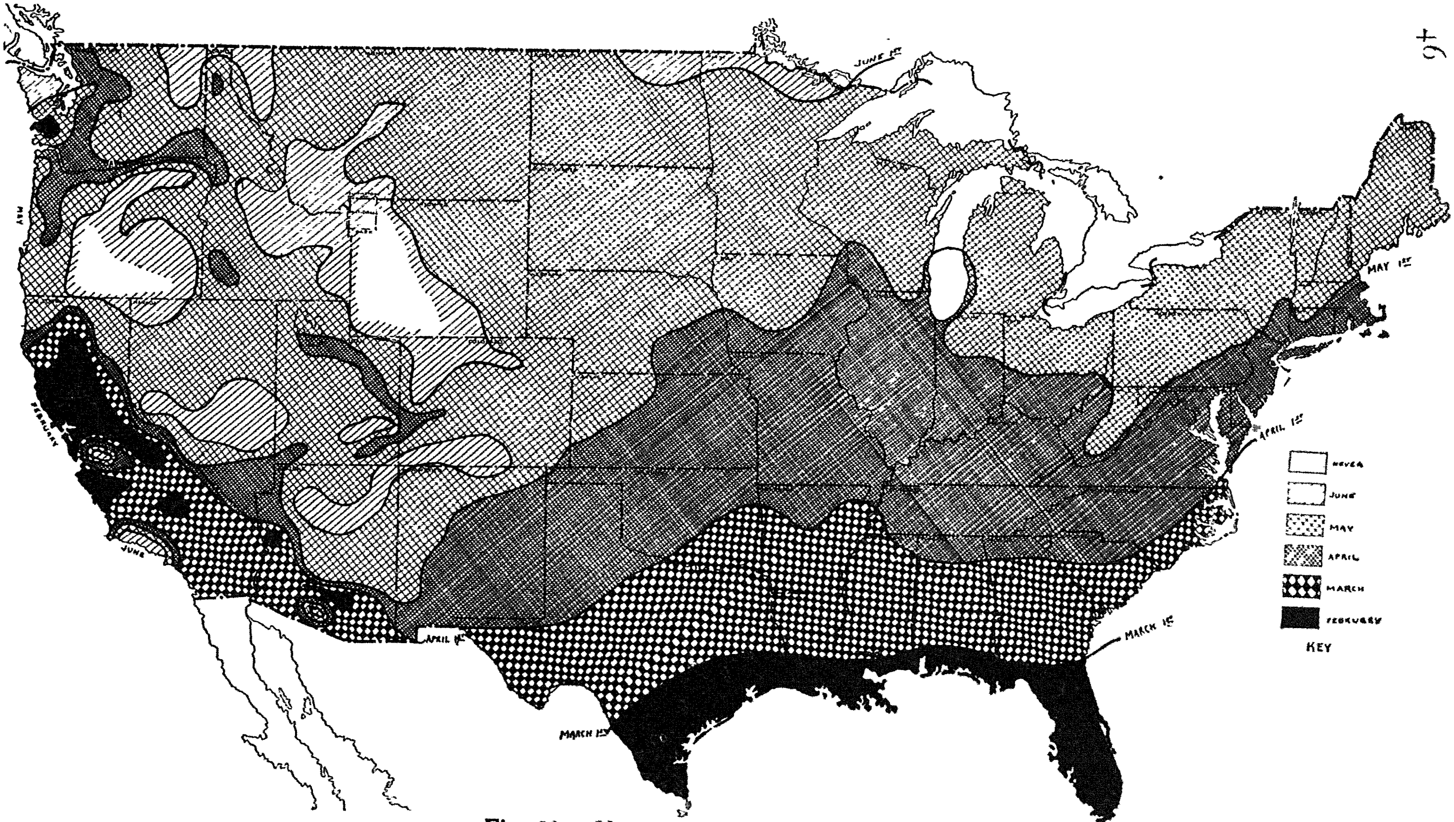


Fig. 39. Classification-chart and Map.
Showing garden-planting times in the United States.

in which the three principles of space, idea, and dynamic relations may be interwoven on a chart. Nor is the number of possible strata or superimpositions necessarily limited to two. We may occasionally need three or four distinct strata, and indeed it is sometimes difficult to tell how many separate strata have been superimposed to effect a single chart.

There are, however, very definite rules which it is well to follow in effecting the combinations. The first is that since two distinct conceptions of relations are being expressed through the same picture, the chart-maker should have clearly in mind their comparative importance, and be prepared to throw the emphasis in his construction upon the more important one. Failure to do this will result in a chart in which the reader's attention is distracted to the unimportant parts, by which it becomes hopelessly confused in trying to follow the important parts. The proper emphasis can be given by using in the superimposed chart, heavier shading or stronger lines which immediately distinguish it at a glance from the lines and shadings of the base-chart. Where color can be used, it should be used judiciously, after studying the effects of different colors to see which gives the right degree of prominence.

The second rule is that since the superimposed chart generally carries the message, the base-chart is usually one with which the reader is already supposed to be familiar. Therefore, when the base-chart, as well as the composite, is unfamiliar to the reader, and particularly when a series of composite charts is to be shown on copies of the same base-chart, it is well to preface the composite charts with a single copy of the base-chart, with which the reader may first become acquainted. Then, when his attention is drawn to the composite, he will be able to use it at once with a full understanding of its significance.

The conception of a basic or underlying pattern, as distinguished from a resulting superadded pattern, is important and stays with us throughout the great majority of mathematical charts which follow. The term "field" will be used for this underlying pattern in these mathematical charts while the upper stratum of curves, bars, and other markings will be known as the "plotting"; and though the choice and construction of the "field" will sometimes be found more important and often more difficult than that of the plotting, yet the field will always be suppressed or submerged by lighter lines and colors, to leave to the plotting its full significance.

PART II. AMOUNT-OF-CHANGE ANALYSIS

Chapter V

STATISTICS

Statistics is a word which has had an unfortunate history throughout its brief existence. To begin with, its pedigree is poor, for it is derived from old Latin words having nothing to do with its present meaning. When it was created a century or two ago, it meant "matters of State," or in dictionary-ese, "matters pertaining to the State." It was then used for those compilations of population, finance, and military strength which rulers liked to have made about their various States. But it has travelled far from that meaning, until today in its proper sense it means any collection of figures and precise numerical information.

The word was rapidly debased, until in common parlance a statistician was one who carried at the tip of his tongue a large assortment of appallingly uninteresting figures on widely irrelevant topics, which he seemed to have memorized from the encyclopedia with the sole purpose of boring us. Now figures are not in themselves necessarily dry and dull—in fact the figures of your bank-account may be very engrossing to you. But figures on uninteresting subjects are a sure cure for insomnia, to all of us. And it goes without saying that if the figures are not of consequence, the chart of these figures will deserve equally little attention. The point is that a chart is as weak as its own data, and a chart-maker must carefully weigh and consider his data before permitting himself the pleasure of illustrating them with a chart.

But a worse charge than mere boredom is often levelled at statistics. It has crystallized into a familiar saying, "figures don't lie, but liars figure." Mark Twain went so far as to remark at one time that there were three kinds of lies—namely, lies, damned lies, and statistics—wicked in the order of their naming. In short, the statistician is sometimes looked upon as one whose acquaintance with figures is so very intimate that

he can readily take liberties with them, abuse them, present them in a false light, and deceive the layman. In this view he is little more than a common trickster, performing legerdemain with numbers, his magical results to be idly wondered at, but not to be trusted. And the moral thereof is clear, that he who would work with figures must be very, very sure that his figures are all correct, both in their computation, and their connotation.

As a matter of fact, surprising as it may seem, we are all in the same boat, a whole nation-full of statisticians, in great or small degree. We all perform mathematical operations, arrange and study figures and precise data. We do it in our accounts, in our reports, our decisions, and sometimes in our sleep. If this be statistics, we are all guilty. And the odium which sometimes attaches to the word must be only superficial, for it does not attach to the practice or the subject-matter for which the word is a symbol. Surely we will not allow ourselves to be daunted by so empty a thing as a symbol or word. Let him who will, retain his shallow prejudice and throw this book aside here and now; and let the rest of us purge ourselves once and for all time of any lingering superstition against "statistics." Let us resolve never again to utter a peep against the word or to be terrified by its use.

There is, it is true, a more precise use of the word "statistics" in which a high degree of proficiency in handling masses of figures is presupposed. In this technical sense, the statistician is one whose ability to digest, compress, or extract significance from a multitude of related numerical data, is highly developed. The science of doing this is called by the queer name of "statistical methods." In point of technical skill it stands somewhere between the science of accounting and the science of higher mathematics. If you have time to dip into it you will find it more interesting than either, because its applications are more varied than the former and more immediately practical than the latter.[1] But for the purpose of chart-making or chart-reading, it is not necessary for you

---

[1] The notable books on the theory of statistics are:

Bowley, A. L., *Elements of Statistics.*

Yule, G. Udney, *An Introductory to the Theory of Statistics.*

Shorter and more elementary texts are:

Kelley, T. L., *Statistical Method.*

King, Willford I., *Elements of Statistical Method.*

Secrist, Horace, *An Introduction to Statistical Methods.*

to have more than the usual grammar-school equipment in mathematics—or as much of it as you have not forgotten.

That the maker of charts must dabble in statistics is obvious; he who would build a house must first examine its foundations. But the chief requisite in examining your statistics is common-sense. You do not need deep mathematical skill wherewith to perform difficult mathematical acrobatics. Common-sense alone, without the aid of calculus or higher algebra, will enable most of us to understand a falling bank-balance, for example. Every school-child knows the meaning of a total, or an average, and all adult persons ought to, without regard for sex, color, or religious persuasion. Apart from such elementary understanding of the meaning of the language, the essential thing in ordinary statistical work is common-sense. If you have it, you can traffic with figures safely, and will often recognize a condition without knowing the technical name or symbol for it, but if you haven't it, all the mathematical skill in the world will merely befuddle you.

Of course, for much work, or for continuous application to statistics, additional mathematical ability is an unquestionable advantage. It is a good thing to know, for instance, that there are several kinds of averages, each with a meaning all its own. It is well to have a nodding acquaintance with dispersion, the word nodding here being used to denote familiarity, not sleepiness. And if you can shake a correlation coefficient by the hand it may help tremendously in a pinch. The man who can write an equation for his profits or his factory conditions, has the edge on the man who has to make lengthy tabulations. But the subject of this book is no more "how to be a mathematician" than "how to get common-sense," and we will therefore drop both problems. We will split fifty-fifty on them, and in proceeding with charting of statistics, assume that the reader is gifted with common-sense, but not versed in higher mathematics. Where special need for certain mathematically technical terms or ideas arises, we will stop on the spot and explain these terms, but we will not build our mathematical bridges until we come to them.

One caution, and one caution only, in this chapter we wish to make so clear that it will remain with the reader throughout the rest of the book. That caution is, do not be afraid to use your common-sense. In this matter an ounce of fore-sight is worth many pounds of hind-sight. The value of this advice

will come to you through long and bitter experience, anyway, for it takes exceptional patience to scrap the results of weeks of research and start all over again at the beginning, just because of a failure to use common-sense beforehand. The most important time to bring your common-sense into play is before you lay pen to paper to take down a single figure. That is the time to ask yourself the all-important question, "What do I want to know?" Unless you can answer that word clearly, positively and concisely, you had better wait until you can, before doing anything else.

"What do I want to know?" Write it down in black and white. Below it write the answer. Stand off and look at that answer as if you were a total stranger, and try to see whether it makes sense. When you finally have the question and its answer so clear that a child can understand, you will be ready to compile and investigate statistics to substitute for the answer. And when your work is finished, and you have boiled down immense quantities of figures and numbers to a simple coherent statement, see if that statement really answers your question. If it does, you have statistics which may be well worth illustrating with charts. If it does not, you had better forego the pleasure of making charts, for it will all probably have to be written off as a waste of time.

You may think this all very simple and easy, but you will find it sometimes immensely difficult, and errors extremely costly. No rules can be laid down, for every case is a matter for individual study and analysis. But the consequences of a mis-step at this stage are grievous. And it is right here that so many amateur statisticians make their first mis-step. They become lost in the zest of hunting up and chasing down all the available related information; they allow themselves to be dragged off the scent of the fox by every jack-rabbit that crosses the trail. Red herring is their meat and in a shorter or longer time they bring home a mountain of "statistics," all of which is "interesting if true," but does not bear directly on the point.

When you find this happening to you, you will be able to recognise it by the bewildered sensation in your solar plexus the first time a friend drops in and asks in a heartless way, "What's the good of it?" And right there is a good time for you to stop—it would have been still better earlier—and ask yourself again, "What *do* I want to know?" Think back once

more to your original position and what you set out to learn. It is never too late to mend. If you find you are on the wrong track, bravely scrap the work you have done, though it hurts cruelly to do so, and strike out again for your goal.

Your goal is a certain piece of information. Statistics are merely the road to that information. The information itself will ultimately reduce to a comparatively simple statement, and if there are any statistics left in that statement, wipe them out by substituting illustrations or charts for them. The graphic method, whether used in the office to facilitate research work, or in the published report to facilitate understanding, should be confined to information which has value. It is therefore an obvious but extremely important rule not to begin a graph until the statistics have been carefully examined and their object or significance brought clearly into mind.

Chapter VI

WORK-SHEETS

What would you think of a factory manager who kept his plant so cluttered up with raw materials, partly finished materials, by-products, and working machinery that his workmen had to climb over each other, and his materials had to be passed from operation to operation by long forward passes skilfully negotiated over ceiling-high piles of obstructions? Yet this is precisely what most of us do in the far more difficult case when our materials, workmen, and implements, are all intangible ideas, expressed by strokes of a pencil on paper. The old copy-book admonition to write clearly and neatly, may often save you from utter confusion and defeat, and will always greatly speed your work. Straight-line computing methods are as important as straight-line factory methods.[1]

Statistical data generally comes in the form of long columns of figures. If it is not already in this shape, you should so arrange it at once. It may be, for example, the reports of your sales in the various States of the country. By listing the States in a column, you can write beside each one the figure of its sales, and your sales will then form a second column. Perhaps you wish to reduce these to per capita sales. In that case, beside the figure for the sales for each State, you can enter the population for each State, thus forming a third column of entries (a second column of figures). The per capita sales, which are merely the ratios between the two columns of figures, can then be entered still further to the right, forming a third column of figures (a fourth column in all).

Frequently the computing is carried through a great many steps, each of which calls for one or more columns of figures.

[1] For an excellent discussion of the principles of tabulation, in addition to the works on statistical methods already referred to (page 49), see Edmund E. Day, "Standardization of the Construction of Statistical Tables," American Statistical Association Quarterly, March, 1920, p. 59.

DISTRIBUTION OF METAL MONEY IN THE WORLD

Approximate Stocks in Chief Countries, Dec. 31, 1918.

(Source: U. S. Statistical Abstract)

	Dollars	Population	$ per cap
TOTAL	10,127,084,000	1,529,179,000	6.62
China	31,358,000	336,042,000	.09
India	175,634,000	315,156,000	.55
Russia	411,600,000	178,905,000	2.30
United States	3,821,363,000	105,015,000	36.39
Germany	543,572,000	67,810,000	8.02
Japan	482,646,000	55,965,900	8.62
Austria Hungary	64,734,000	52,368,000	1.25
Dutch East Indies	49,202,000	47,956,000	1.03
Great Britain	722,561,000	46,089,000	15.68
France	725,449,000	39,700,000	18.28
Italy	249,137,000	36,546,000	6.82
Brazil	43,690,000	26,542,000	1.64
Turkey	-	21,274,000	-
Spain	658,851,000	20,500,000	32.14
Korea	23,889,000	16,913,000	1.41
Mexico	250,000,000	15,502,000	16.13
Egypt	39,376,000	12,566,000	3.13
Siam	41,652,000	8,266,000	5.02
Canada	191,827,000	8,075,000	23.76
Argentina	321,869,000	8,066,000	39.90
Belgium	56,806,000	7,658,000	7.41
Rumania	1,000	7,508,000	-
Netherlands	327,622,000	6,583,000	49.76
South Africa	33,343,000	6,465,000	5.16
Australasia	246,422,000	5,976,000	41.24
Portugal	49,254,000	5,958,000	8.26
Peru	32,691,900	5,800,000	5.63
Sweden	88,856,000	5,713,000	15.55
Colombia	10,768,000	5,071,000	2.12
Morocco, French	24,638,000	5,000,000	4.93
Serbia	33,486,000	4,622,000	7.25
Ceylon	5,776,000	4,262,000	1.36
Switzerland	121,283,000	3,880,000	31.26
Formosa	34,092,000	3,711,000	9.19
Chili	11,363,000	3,641,000	3.12
Finland	-	3,269,000	-
Denmark	52,649,000	2,921,000	18.02
Bolivia	-	2,890,000	-
Venezuela	21,646,000	2,816,000	7.69
Norway	44,911,000	2,509,000	17.89
Haiti	850,000	2,500,000	.34
Guatemala	-	2,119,000	-
Ecuador	4,140,000	2,000,000	2.06
Uruguay	51,094,000	1,346,000	37,96
Salvador	4,398,000	1,268,000	3.46
Paraguay	482,000	1,000,000	.48
Dominican Republic	800,000	725,000	1.10
Straits Settlements	17,263,000	714,000	24.18
Nicaragua	-	704,000	-
Honduras	-	562,000	-
Costa Rica	2,112,000	451,000	4.90
Luxembourg	1,863,000	260,000	7.17
British Honduras	168,000	41,000	4.09

Fig. 40. A Simple Computing Sheet.

The author has, he regrets to say, actually carried one investigation through so many consecutive steps that the resulting columns of figures, when pasted as close together as possible, side by side, without repeating any column, reached completely around the walls of an ordinary room. There is generally no excuse for as much work as this, but it is well to bear in mind that every computing step may call for two or three columns of figures, and that if you are going to carry your figures through many steps, it will pay to have them in uniform columns.

Obviously the arrangement of States or items in the columns should be standard throughout the work, so that any two columns can be readily compared. This makes the order of the items a matter for careful study. In listing the States of the union, for example, you have your choice of three arrangements. In the first place, you can list the States alphabetically, which makes it easy for a stranger to find any particular State at once. Secondly, you can arrange them in the order of their importance (as viewed from the particular stand-point of your problem), which makes it easy for a stranger to focus his attention at once on the important States. But neither of these methods, though widely used, has anything more to recommend it than a certain possible convenience to strangers. For computing at least, the States should be arranged in a logical order, and a logical order is generally one that brings nearby States together, instead of scattering them about the list. The reason for this is that you may find you want to take off sub-totals (totals for East, North, South, and West) to get group-figures for groups of States. And if the States are already arranged or grouped together this is easily done. In fact, it is always well to carry these group-totals, because in checking through to locate errors they save much time.

Having decided to arrange the States logically by territorial groups, you have next to decide how to group them. The census has one grouping, dividing the country into six or seven territories. But this grouping is not the best natural economic one, that is, it does not conform to natural business groupings. The Audit Bureau of Circulations has another, which is, for general business conditions, perhaps the best. But in most cases there will be individual factors which make it desirable to adopt a special arrangement of one's own. Large sales organizations, for example, will already have set up their own sales districts and branch house territories, and in such cases it is best to make the grouping of States conform as much as possible to these.

Of course, not all tabulations are tabulations of States by State figures. These figures illustrate, however, the principles of tabulating. We may have to work with figures, year by year through a series of years, or with month-by-month figures, through one year or more. In these cases, where we are working with divisions of time, the natural and proper way is to place the earliest periods at the top of the list and the

	U.S. Census	Audit Bureau of Circulations	Red Cross
Alabama	NEW ENGLAND	NEW ENGLAND	NEW ENGLAND
Arizona	Maine	Maine	Maine
Arkansas	New Hampshire	New Hampshire	Massachusetts
California	Vermont	Vermont	New Hampshire
Colorado	Massachusetts	Massachusetts	Rhode Island
Connecticut	Rhode Island	Rhode Island	Vermont
Delaware	Connecticut	Connecticut	
Florida			ATLANTIC
Georgia	MIDDLE ATLANTIC	NORTH ATLANTIC	Connecticut
Idaho	New York	New York	New Jersey
Illinois	New Jersey	New Jersey	New York
Indiana	Pennsylvania	Pennsylvania	
Iowa		Delaware	PENN.-DELAWARE
Kansas	EAST NORTH CENTRAL	Maryland	Pennsylvania
Kentucky	Ohio	District of Col.	Delaware
Louisiana	Indiana		
Maine	Illinois	SOUTH EASTERN	POTOMAC
Maryland	Michigan	Virginia	Distr. of Col.
Massachusetts	Wisconsin	North Carolina	Maryland
Michigan		South Carolina	Virginia
Minnesota	WEST NORTH CENTRAL	Georgia	West Virginia
Mississippi	Minnesota	Florida	
Missouri	Iowa		SOUTHERN
Montana	Missouri	SOUTH WESTERN	Florida
Nebraska	North Dakota	Kentucky	Georgia
Nevada	South Dakota	West Virginia	North Carolina
New Hampshire	Nebraska	Tennessee	South Carolina
New Jersey	Kansas	Alabama	Tennessee
New Mexico		Mississippi	
North Carolina	SOUTH ATLANTIC	Louisiana	LAKE
North Dakota	Delaware	Texas	Indiana
Ohio	Maryland	Oklahoma	Kentucky
Oklahoma	District of Columbia	Arkansas	Ohio
Oregon	Virginia		
Pennsylvania	West Virginia	MIDDLE STATES	CENTRAL
Rhode Island	North Carolina	Ohio	Illinois
South Carolina	South Carolina	Indiana	Iowa
South Dakota	Georgia	Illinois	Michigan
Tennessee	Florida	Michigan	Nebraska
Texas		Wisconsin	Wisconsin
Utah	EAST SOUTH CENTRAL	Minnesota	
Vermont	Kentucky	Iowa	GULF
Virginia	Tennessee	Missouri	Alabama
Washington	Alabama	North Dakota	Louisiana
West Virginia	Mississippi	South Dakota	Mississippi
Wisconsin		Nebraska	
Wyoming	WEST SOUTH CENTRAL	Kansas	NORTHERN
	Arkansas		Minnesota
	Louisiana	WESTERN STATES	Montana
	Oklahoma	Montana	North Dakota
	Texas	Wyoming	South Dakota
		Colorado	
	MOUNTAIN	New Mexico	SOUTHWESTERN
	Montana	Arizona	Arkansas
	Idaho	Utah	Kansas
	Wyoming	Nevada	Missouri
	Colorado	Idaho	Oklahoma
	New Mexico	Washington	Texas
	Arizona	Oregon	
	Utah	California	MOUNTAIN
	Nevada		Colorado
			New Mexico
	PACIFIC		Utah
	Washington		Wyoming
	Oregon		
	California		NORTHWESTERN
			Idaho
			Oregon
			Washington
			PACIFIC
			Arizona
			California
			Nevada

Fig. 41. Various Geographic Groupings of the States.

latest at the bottom, with possibly space for quarterly, half-yearly, or five-yearly totals, which ever may be desired. Still other types of items may occur, such as in a list of the various departments of a plant, the various salesmen of a selling organization, or the various products, and so on.

The point is that whatever the items be with which we work, they should be arranged carefully at the outset, so that it will not be necessary to alter their arrangement later. They should be placed in a column if possible, so that the computing and tabulating can be made in parallel columns beside them. If the list is long, it should be broken up by blank spaces at convenient intervals, or better still, the items should be grouped together, for which sub-totals may be required, and blank spaces should be inserted between the groups for these sub-totals or part totals. But by all means, try to get the whole list on a single page, even at the cost of pasting additional sheets at the bottom, for the work will progress much faster on one large sheet which is complete than on several small sheets which are not complete.

If much work is going to be done, it pays—and pays well—to have the printer rule up some sheets with the list of items printed at the edge of the sheet and the lines ruled in where they will be useful, horizontally from each item. It is a small matter, but worth noting, that the lines should be regular typewriter distance apart, so that should you wish to have figures typed on these sheets the typist can work rapidly and neatly. More important still, where adding machines are

	(Tons)
NEW ENGLAND	
N. Hampshire, Mass., & Conn.	10,300
MIDDLE ATLANTIC	
New York & New Jersey	2,600,000
Pennsylvania	14,000,000
SOUTH-EASTERN	
Maryland	524,000
Virginia	429,000
Alabama	2,390,000
Tennessee	283,000
Kentucky	772,000
NORTH CENTRAL	
Ohio	8,530,000
Indiana & Michigan	2,940,000
Illinois	3,280,000
WESTERN	
Iowa, Missouri, Colo., Mont., & Ore.	465,000

PIG-IRON PRODUCTION
United States
1920

Fig. 42. An Incomplete State-list Geographically Arranged.

used, is to get the lines spaced in the same way that the adding machine prints the tape, so that figures do not need to be copied from the tape, but the tape can be pasted right on the sheet. In wide carriage machines, the tape can be dispensed with, and the figures printed directly by the adding machine on the sheets. These are devices to speed up the work, which are trivial in themselves but very important in their results, and, with a careful eye to your equipment, you will soon hit upon the most useful forms for doing your computing, standardise them, and call in the printer to prepare a number of blanks.[2]

We have here considered carefully only the matter of the items which flank the left hand edge of your work-sheet forms. Technically, these items are called "the stubs," indicating that they are the labels attached to each horizontal line, or row of figures in the columns or column of the table. The whole sheet, filled with figures in orderly arrangement, is called a "table" or tabulation. The vertical lines of figures are called "columns" or sometimes "arrays," while the horizontal lines of figures, that is, the sets of figures beside each stub, are called "rows" or lines. And at the top of each column of figures, the label which describes the column in the same way that the stubs describe the rows, is called a "heading" or "caption."

A whole chapter could be written about captions, or column headings. It is a fine art to make them at once clear and brief, and to arrange them so logically that the thought moves easily from one heading to another. In most cases of several columns, two or more captions can be bracketed together by a third common group-caption. In this way, the headings often take on the form of miniature classification-charts. The best practice is to box in the headings carefully, so that they will be clearly understood. In work sheets they should be arranged

[2] The position of the total (or sum) in a tabulation is an important matter. There are two possible positions: first, at the beginning or top of the table; second, at the end or bottom of the table. The first is the statistical position; it is correct for a published table, as it places the most important item, the total, or whole, first before the reader's eye. When parts are themselves further subdivided, their totals should also be placed before their parts. The details, or parts, can be in every case further indented than the totals. This practice should be adhered to in charts and in published or recorded tables.

The second position, at the end of the table, is the accounting position. It is correct for all cases in which the work of summing up the parts must be frequently undertaken. Needless to say, it is essential for forms to be used in adding and listing machines, and is advisable in general for work-sheets.

ILLITERACY IN THE UNITED STATES
Illiterate percentage of each class (by age, sex, race) of the population for each group of states, 1920
(Source:- U. S. Census)

	Youth (Aged 16-21 years)	Adult (Over 21 years) Male.	Female.	Total population over 10 years of age. Native white	Foreign born white	Negro	Total
TOTAL U. S.	5.3	7.0	7.3	2.0	13.1	22.9	6.0
New England	1.1	6.0	6.2	0.7	14.0	7.1	4.9
Middle Atlantic	0.8	5.9	6.6	0.6	15.7	5.0	4.9
East North Central	0.5	3.7	3.6	0.9	10.8	7.3	2.9
West North Central	0.6	2.5	2.5	0.9	6.4	10.5	2.0
South Atlantic	7.9	14.0	15.9	5.1	12.8	25.2	11.5
East South Central	7.6	15.7	15.2	6.4	9.1	27.9	12.7
West South Central	7.3	12.3	12.1	4.1	29.9	25.3	10.0
Mountain	3.7	5.4	6.8	2.0	12.7	5.3	5.2
Pacific	1.2	3.3	3.0	0.4	8.6	4.6	2.7

Fig. 43. Classified Headings to the Columns.

in the order in which they will be computed, so that the work of computing moves as much as possible to the right, always preferably deriving each column from the immediately preceding columns. Where two or more columns, which are in themselves the results of several columns, must be combined, each can be left at the right-hand edge of separate computing sheets and then by cutting off the remaining paper, or by folding it back, you can lay the sheets one over the other so that the columns to be worked over will appear side by side.

One of the cardinal rules for computing tabulations—and it applies only in lesser degree to final tabulations ready for publication, presentation, or study—is that every column head should include a number or letter identifying the column. This rule has even been extended by some authorities to the stubs as well. The advantage of the number or letter is that it makes reference to the particular column very easy, either in texts, notes, conversation, or formulae.

In addition to a symbol for the column, you should also have in your column-head, a note explaining the source of the figures in the column. This note can either be the name of the authority from which the figures were copied, or it can be the formula by which the figures were computed from other columns. It will save you much trouble in correcting the errors of computing clerks, and much time later on when you come to refer to the sheets and do not remember the various

CLASSIFICATION BY AGE & SEX / CLASSIFICATION BY RACE AND BIRTH		THE OLDER GENERATION (Population born before 1899)										THE YOUNGER GENERATION (Born between 1899 and 1910)			Relative Illiteracy of the two generations
		MEN			WOMEN			Relative Illiteracy of Men and Women	TOTAL			TOTAL			
		Total	Illiterate		Total	Illiterate			Total	Illiterate		Total	Illiterate		
		Number	Number	Percent	Number	Number	Percent	(Men = 100)	Number	Number	Percent	Number	Number	Percent	(Younger = 100)
		Copied	Copied	B/A	Copied	Copied	E/D	F/C	A+B	B+E	J/H	Copied	Copied	M/L	K/N
		A	B	C	D	E	F	G	H	J	K	L	M	N	P
WHITE	Native White	21,513,948	563,546	2.62	21,100,793	477,121	2.26	87	42,614,741	1,040,667	2.42	18,247,122	201,905	1.10	220
	Foreign born White	6,928,452	840,063	12.12	5,570,268	867,082	15.55	128	12,498,720	1,707,145	13 68	939,136	56,595	5.56	241
	Total	28,442,390	1,403,609	4.94	26,671,061	1,344,203	5.04	102	55,113,461	2,747,812	4.99	19,246,288	258,500	1.34	375
COLORED	Negro	2,792,006	748,229	26.8	2,730,469	764,758	27.9	104	5,522,475	1,512,997	27.4	2,530,750	329,174	13.0	211
	Other Races	168,974	40,530	23.9	81,620	31,782	39.0	163	250,594	72,312	28.8	75,747	11,120	14.7	196
	Total	2,960,980	768,759	26.6	2,812,089	795,540	28.4	107	5,773,069	1,565,299	27.4	2,606,437	340 294	13.0	211
TOTAL		31,403,370	2,192,368	6.97	29,483,150	2,140,743	7.26	104	60,886,530	4,333,111	7.11	21,852,785	598,794	2.75	259

ILLITERACY BY AGE, SEX, AND RACE
United States
1920
(Arranged from United States Census)

Fig. 44. Column Symbols, Formulae, and Classified Stubs and Captions.

steps clearly. For the clerks who are doing the computing, this note is a standing bill of instructions. And the time will

SAVINGS BANK STATISTICS
Number of savings banks and savings bank depositors; total, average, and percapita deposits; and ratios between banks, depositors, and population as specified below.
United States, 1820-1920
(Arranged from U. S. Statistical Abstract)

Year	Banks	Deposits	Depositors	Average deposit	Population	Per-capita deposit	Depositors per Pop.	Depositors per Bank	Pop. per bank
	Number	Dollars	Number	Dollars	Number	Dollars	Percent	Number	Number
	Copy	Copy	Copy	B/C	Copy	B/E	100 C/E	C/A	E/A
	A	B	C	D	E	F	G	H	J
1820	..	1,138,576	8,635	132	9,638,453	.12	.09	..	..
1825	15	2,537,082	16,931	150	11,150,000	.23	.15	1,130	744,000
1830	36	6,973,304	38,035	183	12,866,020	.53	.30	1,056	357,000
1835	52	10,613,726	60,058	172	14,710,000	.72	.41	1,155	283,000
1840	61	14,051,520	78,701	179	17,069,453	.82	.46	1,290	280,000
1845	70	24,506,677	145,206	169	19,970,000	1.23	.73	2,075	285,000
1850	108	43,431,130	251,354	173	23,191,876	1.87	1.08	2,325	215,000
1855	215	84,290,076	431,602	195	27,256,000	3.09	1.58	2,050	126,700
1860	278	149,277,504	693,870	205	31,443,321	4.76	2.20	2,495	113,100
1865	317	242,619,382	980,844	247	34,748,000	6.99	2.83	3,095	109,600
1870	517	549,874,358	1,630,846	337	38,558,371	14.26	4.24	3,160	74,500
1875	771	924,037,304	2,359,864	408	43,951,000	21.00	5.37	3,060	56,900
1880	629	819,106,973	2,335,582	366	50,155,783	16.30	4.66	3,710	79,800
1885	646	1,095,172,147	3,071,495	356	56,148,000	19.50	5.47	4,750	87,000
1890	921	1,524,844,506	4,258,893	359	63,056,438	24.18	6.59	4,630	68,400
1895	1,017	1,810,597,023	4,875,519	372	69,579,868	26.04	7.01	4,790	68,400
1900	1,002	2,449,547,885	6,107,083	401	76,129,408	32.18	8.02	6,100	76,100
1905	1,237	3,261,236,119	7,696,229	424	84,219,378	38.80	9.12	6,220	68,100
1910	1,759	4,070,486,246	9,142,908	445	92,267,080	44.20	9.91	5,200	52,500
1915	2,159	4,997,706,013	11,285,755	443	99,342,626	50.40	11.35	5,220	46,000
1920	1,707	6,536,470,000	11,437,556	571	106,418,175	61.20	10.73	6,700	62,300

Fig. 45. Column Symbols and Computing Instructions.

come when you will map out your work in this way, merely filling out stubs and the column-headings yourself, and leaving the entire computing task to clerks.

It is also just a question of time before you will come to look upon these tabulated figures as so many various descriptions or phases of the original set of stubs. The column-headings tell you the type of description or phase, but in the end the stubs are the basis, and the figures are derived from them. Mathematicians have a word "function" for such relations. Using that word, we would say that the tabulated

figures are functions of the stub-figures or items, meaning that their values are derived therefrom. Glancing across the various lines, you will see that all the figures on the same line with a stub are but various and varying functions of that stub, the symbol at the top of the column identifying the function and the caption describing it.

Glancing down any column, you will see that the figures change from item to item. They form a series of varying values. Each column contains the various values of the function described by the caption head. Each column can be looked upon as a series of figures which are the readings or values of the item, described by the column head. And the point is that while the stubs, and captions, were independently arranged by yourself, the figures in each column are derived from or attributed to them and so are dependent upon the stubs and captions. In short, while the stub is an "independent variable," the series of figures in the other columns are all "dependent variables" with regard to the stub.

Sometimes the independent variable is called the "x-variable" and the dependent variable the "y-variable." In that sense the values of "y" will all depend upon the values or meanings of "x." We might go so far as to label the column of stubs "x" and the captions "y," being various kinds of "y" variables. Think of the first one as "y," the second column as "y_1," the third as "y_2," the fourth as "y_3," and so on, and it will always be clear to you that the stub is the independent or x-variable and the other columns are the dependent or y-variables. A second table on an entirely different aspect of the same items or stubs, could be called a "z-variable," meaning that while it was also a dependent, it was distinct from the first set of dependents. As a matter of fact, this is all relative, for at some time you may treat one of the columns of figures as an independent variable and the stub as its dependent. But it is useful to begin thinking of your tabulation as having both independent and dependent variables contained in it.

Chapter VII

CO-ORDINATES

The better to demonstrate their simplicity, we have discussed in the very first chapter those few technical terms which will be essential to the student in this elementary work. In that chapter the analogy is drawn between the checker-board arrangement of streets in certain American cities and the criss-cross rulings of most chart paper. In this chapter this particular type of ruling will be carefully re-considered for the benefit of those whose zeal in the subject has led them to skip the first chapter, and a few other forms of ruling will also be touched upon with their relations to this fundamental system. The chapter will not be interesting reading, but it deserves close study in order that the remainder of the subject may be clearly understood.

Station yourself in an open field where you can, without difficulty, move in any direction—even upward with the aid of an aeroplane or downward with the aid of a good fast shovel. At the point where you stand, drive a stake into the ground and consider it your base of operations for all other points on, above, or below the field, that is, the point from which you can measure their distances. We will call it the "point of reference" or "point of origin" and mark it "O," which according to your taste may either stand for the word "origin" or for its zero distance from the origin.

Facing in any direction at this point walk forward a distance, let us say, of five steps, in a straight line. It is obvious that you could walk forward indefinitely in this direction, always reaching a greater distance from the starting-point "O,"

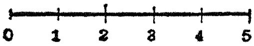


Fig. 46.

and that you could always measure this distance by stakes at each foot-step numbered "1," "2," "3" and so on successively

as you pass them. You could then instantly locate any point along your path by the number of the stake. Points midway between full steps can be given fractional values. Such points would be definitely and unmistakably identified if, in addition to telling their distance from the origin, you also tell the direction in which you had walked when you left the stake to measure them. Calling this direction "*x*," you would specify the points completely by calling them "*x*, 1," "*x*, 2," "*x*, 3," and so on.

Having walked forward, however, only five steps, you would reach the point "*x*, 5." Suppose here you stop and begin to retrace your steps, walking backwards. You would notice that now instead of the distances from "O" increasing, they decrease as you walk backwards, until after five backward steps you are again at the zero-point. But continue to walk backward and you will have again the phenomenon of increasing distances from this point as your steps increase. In other words, along the same straight line, there are two sets of distances mirroring each other at the zero-point. If you walk backward ten steps all told, from the point "*x*, 5," you come to another spot which is also five steps from the origin and in the same straight line. There must be some way to distinguish the two and to distinguish all the other pairs of distances in this straight line which we have called "*x*." Suppose that for this purpose every distance reached by walking forward from the origin be called positive and every distance reached by

Fig. 47.

walking backward from the origin be called negative. Leaving the first set of stakes as before marked "*x*, 1" (or "*x*, +1"), "*x*, 2" (or "*x*, +2") and so on, you can similarly mark the stakes at your backward steps, from the origin, by numbering them "*x*, −1," "*x*, −2," "*x*, −3" and so on. In this way you will easily identify every point along the straight line, both on one side of the origin (in the "*x*" direction) and on the other side (in the "−*x*" direction).

Here you have set up measurement in one direction, that is, along one "dimension." There is a technical name for this "direction" or "dimension," which it will pay you to learn, namely "axis." In fact, we can already speak of it as the

"x-axis," or axis of "x" measurements. But leaving aside technical terms, your common-sense will tell you that you have set up the only kind of measurement possible in one dimension, namely "linear measurement." In this particular case, your foot-step—one half of your pacing distance—is your "unit of measurement." Any other unit could have been taken. You might have laid a certain stick down repeatedly and marked off the number of times it could be laid end over end. If you had a number of sticks of the same length you could lay them end to end and leave them lying to form one long pole or rod cut into equal parts. But whatever the length of your unit of measurement may be, you will, by numbering the units successively in both positive and negative directions, have graduated or "calibrated" that one long imaginary rod with a "scale"—the scale or calibrations being the numbers or countings from the zero-point in both directions.

It will occur to you, however, that while you have an excellent system for locating points in that one line along which this imaginary rod lies, that is, along which you walked, you have no means of identifying points elsewhere in the field. Suppose, therefore, you return to the "origin" with a short actual rod which is long enough to reach from point "x, 5" to point "$x, -5$" and lay it actually between those two points. (You can extend this rod in your imagination indefinitely, in both directions, but a rod of ten steps length is handier to carry about than an indefinitely long one.) At every stake mark the corresponding point on your rod, with the same numbers, so that you can dispense with the stakes entirely, and with the rod in this position you can still locate points along the "x" direction from the origin, either positively or negatively, at once.

Now stand at the origin with the positive markings on the rod to your right and its negative markings on your left hand, and you will find yourself facing in a direction at right angles to your first line of walk. Strike out in this new direction. As you walk you can again keep track of the distance by counting foot-steps, and you will again in this new direction find both positive and negative values which mirror each other. You can find positive values by walking forward and negative ones by walking backward from the origin. This new direction, lying at right-angles to the original "x" direction, you can call the "y" direction, and specify points along it as "y, 1,"

"y, 2," "y, 3" and "y, -1", "y, -2," "y, -3," and so on. And so you will have set up a second scale at right angles to

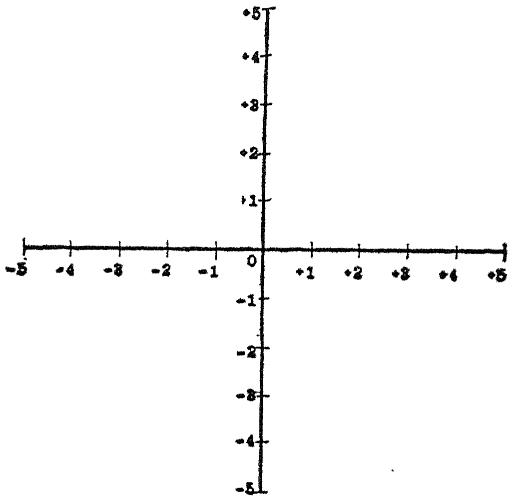


Fig. 48.

your first or "x-scale," by means of which you can identify points along the "y" direction from your origin. The unit of measurement may or may not be the same as the unit used in the first measurement, but of course whatever unit you adopt in the "y-scale" must be adhered to as closely therein as was the "x-unit" along the "x-axis." Being in a new direction, the distances have no relation to those in the old direction, but it is obvious that along either direction the units therein must be uniform. And since we called the first direction the "x-axis," we may call this new one, at right angles to it, the "y-axis."

Now if you will pick up the rod, lying in the "x-direction" and carry it with you as you walk in the "y-direction," without swinging it or moving one end faster than the other, but taking care to hold it rigidly as you walk, you will see that every point along the rod, marking distances along the "x-axis" describes a straight line parallel to the "y-axis" in which you walk. Here you will have a means for identifying any point on the field. You need but to carry the rod or x-axis out along the y-axis until the rod crosses the point you wish to identify. Then note the point on the rod with which it coincides, such as "x, 3," and the point on the y-axis to which

you carried the rod, such as "y, 6." You will define the point as being "x,3; y, 6," and you will search in vain for any other

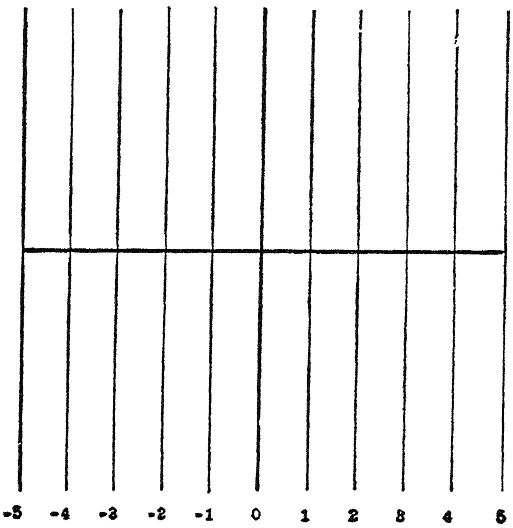


Fig. 49.

point which can be similarly described. Three other points you will find which have the same numbers, but not the same signs; these points will be "x, 3; y, −6, "x, −3; y, 6" and "x, −3; y, −6." You will find that your two axes cut the field into four quarters, in each of which all points have the same combination of signs.

If you do not wish to carry the rod back and forth along the y-axis, or a similarly marked rod in the y-axis back and forth along the x-axis, you can lay a series of rods parallel to each other and perpendicular to the x-axis, crossing that axis at each point marked off on it, and another set of perpendiculars along the y-axis so that your entire field is crossed by these measuring rods in two directions. Thereafter you can locate any point on the field by walking out either axis and then turning at the right distance and following the perpendicular there, parallel to the other axis.[1]

You now have at your disposal a means for identifying any point or any number of points upon a given field, or in precise

[1] The co-ordinate axes need not be at right angles to each other, they may be drawn at any other angles desired; in the latter case, they may be thought of as really at right angles, but seen from a side rather than a direct view.

language, in a plane surface. A plane has, as geometricians say, two dimensions, commonly called length and breadth.

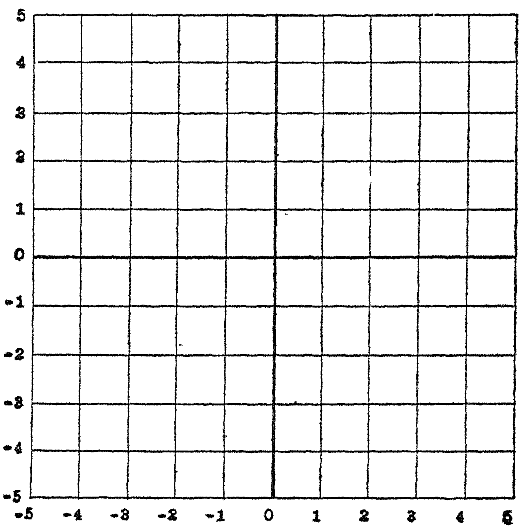


Fig. 50. Field with Equal Scales.

For these you have laid down two straight lines, or axes, at right angles to each other. One of these you have called the "x-axis," the other the "y-axis." Both have been calibrated or measured off and scales attached. And with this simple mechanism you can take measurements of various objects in your field and record locations so precisely that others can, by your records, be led to the very same objects, or can discover without possibility of doubt, the exact spots upon which the objects had been placed when you measured them. In

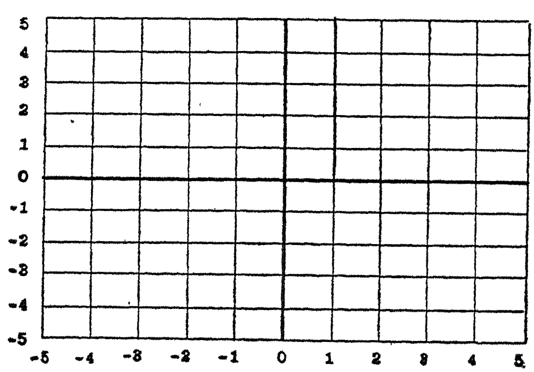


Fig. 51. Scales of Axes Unequal.

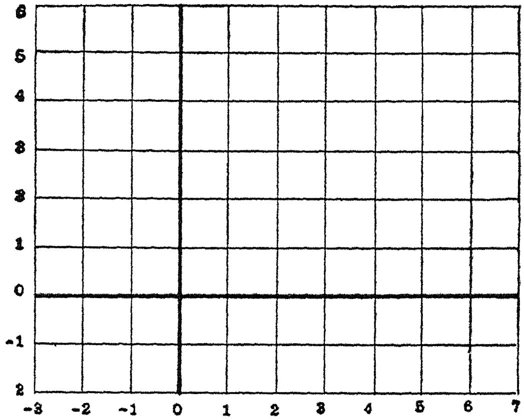
short you have the means for identifying any point upon a plane surface.

It is a fortunate coincidence that the paper upon which we ordinarily write is flat and its surface can be considered a plane surface. For this enables us to apply a reverse English to our measuring device, and use it in making precise illustrations of such fields as we have measured. Now instead of being given a field with objects and being required to set up the measuring device in order to ascertain the location of the objects, we will be given paper already ruled off with the measuring device, and will be required to place thereon indications of the objects.

The paper ruled off in this way is called "co-ordinate paper" for reasons which will presently appear. Its rulings represent a series of parallel lines laid crosswise over another series of parallels, and we are at liberty to select any two intersecting lines for our axes and mark off our scales or measurements along these lines as large or as small as may suit us. And you will find that more than half the charts you ordinarily encounter will be constructed in this way. The horizontal lines are called abscissae, the vertical ones ordinates.[2] Taken together, these co-ordinate rulings form what is, in chart-making, technically called the "field" of the chart, being the background upon which the distinctive portion of the chart is superimposed. And as will be remembered, from the chapter on superimpositions, the basic portion or field should be as unobtrusive as possible, that the important features may receive more attention. The field should always be drawn lightly, with thin lines, and with no more co-ordinates ruled in than are necessary to afford the chart-reader ease of comparison. If possible, the field should be in green or grey ink, as this further submerges it.

The origin or zero point we have so far taken within the field, so that the field is cut into four quadrants in which the

[2] In consequence of that quaint genius for unnecessary trouble sometimes exhibited, a very serious discussion has occasionally arisen as to whether the abscissae should be ruled upon the paper with their upper or lower edges upon the exact positions which the lines signify. The idea of those who favor the upper edge appears to be that the abscissae are shelves upon or above which plotted points rest, when permitted to do so. The idea of those who favor the lower edge is not so cogent. Curiously enough, similar debate has never raged around the position of the ordinates. As a matter of fact, few charts are so precisely drawn or so finely adjusted that the thickness of the ruled line is material, and in every case, the obvious place for all ruled lines is a centered position, that is, one in which their two edges are equidistant from the precise desired positions of the imaginary lines they represent.



ig. 52. Origin of Chart near One Edge of Field.

values of points mirror and repeat themselves with different plus and minus signs. But in practice a large proportion of our measurements or data present no negative values, at least

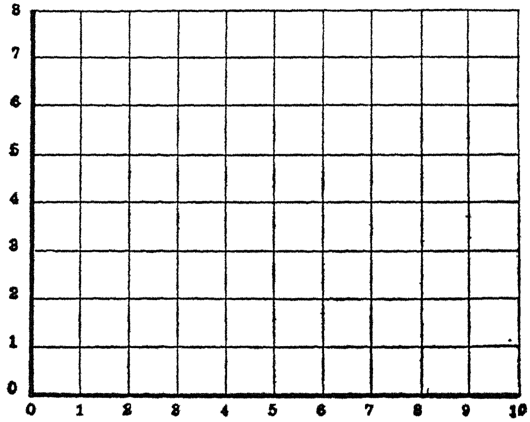


Fig. 53. Origin in Corner of Field.

along the x-dimension. The right-hand upper quadrant, or at most the two right-hand quadrants are then sufficient, and we can omit the remaining quadrants entirely. As a result, the ordinary chart on these co-ordinate rulings shows the value of zero or its origin-point along its left hand edge and generally in the lower left-hand corner. The method is still the same, but a portion of the field is merely being omitted, because useless. In fact, we can even go further and begin the chart

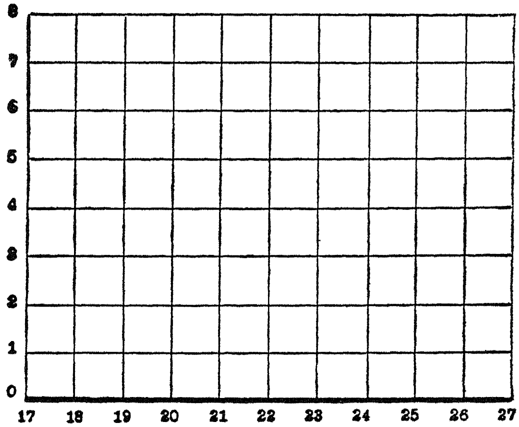


Fig. 54. Origin Not Shown in Chart.

out to the right of the origin, omitting the origin itself when that also is not to be used.

This system of parallel lines to two axes, which are themselves at right angles to each other, belongs to what is called the

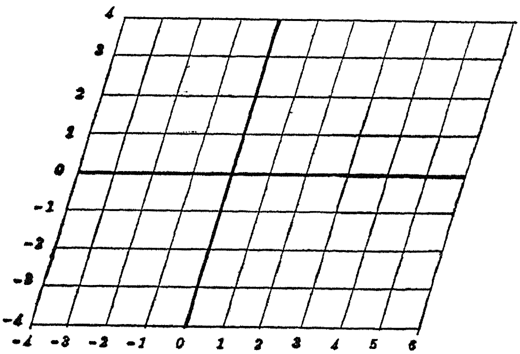


Fig. 55. Field with Co-ordinates Not Perpendicular.

Cartesian system of co-ordinates. That system need not stop with plane two-dimensional surfaces but can be extended to cover three-dimensional space, by the very simple expedient of perpendiculars erected at the intersections of the parallels. You may think of this as lifting your entire net-work of rods up over your head as you stand on the "origin" in your field, or forcing it down into the earth below you; calling the vertical line through "O," the "*z*-axis" and measuring upward distances positively and downward distances negatively. But because three dimensions are not easily represented on a flat piece of

paper, you will find few charts which have at the same time three axes.

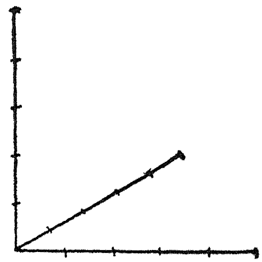


Fig. 56. The Three Axes of Three-dimensional System of Perpendicular Co-ordinates.

Students of trigonometry will recall another method of measuring. They will say, standing at your point of origin with the "x-axis" rod in your hands, do not walk forward, but turn slowly around. The points on the rod will then describe a circular movement about you, and you can locate any point

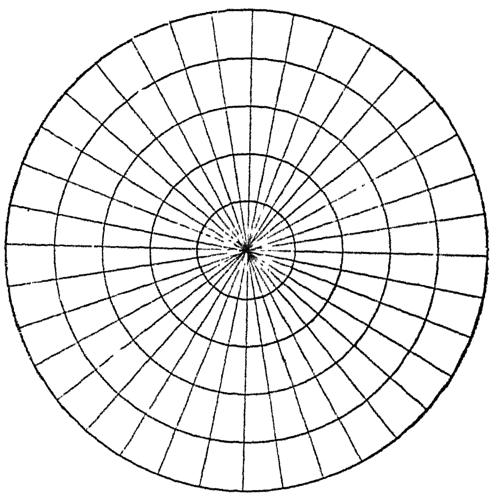


Fig. 57. Polar Co-ordinates.

in the field by noting the distance on the rod and the angle through which it has been turned. Indeed, you will find many a chart which is built on this principle. Later on when we come to some examples of it, we will have to explain its peculiar qualities. For the present it is enough to say that the method of "polar co-ordinates" can be used.

The great mass of chart work is built along plain co-ordinates. The simplest form of all uses but one axis, and consists only of a straight line; the commonest form uses two axes, and consists of square or rectangular outlines. Occasionally we have use for three axes, using three dimensions, and we must not forget also, the circular two-dimensional device with polar co-ordinates.

All this is so extremely simple that we hesitate to dwell so long upon it. But unfamiliar names are great bug-bears; let the idea be as simple as pie, technicians will come along and give it a long high-sounding name, generally created on the spur of the moment by themselves out of ancient Latin or Greek dictionaries, but which even the old Romans and Athenians themselves would have been unable to understand. We must not let them fool us with these long empty names. On the contrary, when we see "co-ordinates" we will know that it means nothing more than criss-cross lines. We will remember the criss-cross checker-board arrangement of roads in American cities, and unless we hail from Boston, we will think of "street" when we see the word "abscissa" and we will think of "avenue" when we see the word "ordinate." The "x-axis" is "Main Street" from which on both sides the other streets (abscissae) are counted; the "y-axis" is "Main Avenue" from which on both sides the avenues (ordinates) are counted.[3] The point is that under no consideration will we let ourselves get disturbed by the unfamiliar names, and in a little while we shall be able to swing these words around with the best of them.

[3] "Main Avenue" (the "y-axis") is crossed or cut by the streets (abscissae—Latin for cut-offs) and "Main Street" (the "x-axis") is crossed or cut by avenues (ordinates—their Latin failed them here—ordinates means arranged in order). More than that, if the streets run east and west, and the avenues north and south, then positive x is east along a street from Main Avenue and negative x is west; positive y is north along an avenue from Main Street and negative y is south. Upper Central Park West, for example, is negative x from Fifth (Main) Avenue and positive y from Battery Park.

Chapter VIII

DIMENSIONS AND VARIABLES

While it is deemed essential for artists to understand the nature of crayons, brushes, and like materials for their work, yet we often observe that they make a great to-do over their study of human anatomy, landscape scenery, and the subject-matter of their work in general. In the same way we who would illustrate mathematical facts must, it is true, mind our "p's" and "q's"—which is to say, our "abscissas and ordinates," but these are merely the tools with which we work, and most of our attention should be directed to our subject-matter—the data or statistics we wish to illustrate. The closer is our analysis and understanding of the nature of this subject-matter, the clearer and more accurate will be our graph or illustration of it.

Because this book is largely a manual on the craftsmanship of charting, and will therefore be mostly devoted to the examination of the various types and kinds of charts with which you can illustrate statistical facts, we here take one last occasion, before the curtain rises on the charts themselves, to appeal to the reader, on behalf of his own common-sense, and urge him to review carefully his analysis of his statistics, before laying ruler to chart.

We make this appeal at this time when the reader has finished his statistical work and is addressing himself to its charting. This is the time when, once more, he must stand off and scrutinize the statistics which he has compiled in the face of such great difficulties—the statistics which are as dear to him as his own thought-children—and scrutinize them with cold, calculating, dispassionate eyes. Do not make the mistake of plunging into the charting-work direct from the statistical work, but once more carefully weigh those statistics against the original question, "What do I want to know?"

The reason for this is two-fold. In the first place, this final examination will give you valuable suggestions as to what part of the statistics the chart should emphasize. It will enable you to eliminate from your chart entirely what is not significant to your inquiry, though the same may be necessarily retained in the table for reference. It will enable you to focus the chart upon the essential information and present it to others in precisely the light or relation in which you wish it seen.[1] In short, it will take the "straw" out of your graph and leave only the wheat.

In the second place, you need this last-minute scrutiny of your figures, or some similar analysis when your results have come clearly into view, to determine how you will chart the statistics. Though it will not always decide the precise form of chart for you, it will settle the fundamental principles of that chart, and you can easily modify details later. The remainder of this book will be devoted to the details of the various charts, but here and now let us attack the fundamental principles.

You have seen in the last chapter that measurement of space is based upon linear distance in each of its dimensions. A one-dimensional space would require but one axis or direction for measurements. A two-dimensional space would require measurement in two axes or directions. A three-dimensional space requires measurement in three axes or directions. In short, there must be as many axes for measurement as there are dimensions.

Now in precisely the same way, your statistical data can be considered as having one, two, or more dimensions. If there is but one dimension in your data, you should use but one axis or direction on the chart-paper to picture it. If there are two dimensions in your data you can use several one-dimension charts or one two-dimension chart. If there are three dimensions in your data, you can either present them two at a time in several two-dimension charts, or try your hand at one three-dimensional chart—a more complicated form.

---

[1] It is not intended here to suggest that it is possible or desirable to practice deception, with any correctly-made chart, but only that the chart-maker can bring out various aspects of the truth about his data with greater or less clearness, by his choice of charting method.

How can you tell how many dimensions there are in your data? You can tell by the form which your data takes when you have neatly tabulated it. You must allow one more dimension to your chart than the figures to be plotted actually require for correct tabulation. That's the rule. For a figure itself must be considered as having one dimension. (Imagine each figure as reaching up perpendicularly off of the paper and this will become clear to you.)

This rule can therefore be stated in another way. For single figures use one dimension or axis on your chart. For a series of figures, arranged either downward in a column or across in a row, use two dimensions. For a series of series of figures, comprising a row or line of columns side by side, use three dimensions. These last cases are not frequent, and are limited to occasions where you could have presented the figures in each *row* in a two-dimension chart, or the figures in each *column* in a two-dimension chart, but wish instead to show the figures by *columns and rows* simultaneously, requiring, therefore, a three-dimension chart.

It is always possible to present several charts in one. For instance by using a number of one-dimension charts, you may adhere to the one dimensional principle, but show several at the same time. A series of horizontal bars, for instance, is an illustration of this. (But when the series is arranged with care as to the downward axis also, these many one-dimension charts can sometimes be made into one two-dimension chart.) Or you may use a number of two-dimension charts, thereby adhering to the two-dimensional principle, but showing several at one time. A number of curves on the same chart-field illustrate this. However, where the base-lines are differentiated with care, these can often be made into one three-dimensional chart. Charts into which several single charts have been compressed are called multiple charts.

This brings us to an important exception or modification of the simple rule for dimensions. For it is necessary that you distinguish between variables and other functions. A table which includes two or more functions which are not variables is really not a simple table but a multiple table, into which several simple tables have been combined. Thus, if the stubs of your tabulation (that is, the items at the left of each line) form a variable, they should be counted as requiring a dimension on the chart. But if they are not values of a mathe-

matical variable, they do not add to the dimensions of the chart but merely form a list of the number of charts which may or may not be compressed into one multiple chart. The same considerations hold true of the "column-headings" or "captions" of the tabulation. The stubs or headings constitute variables, when their arrangement is fixed by their mathematical sequence, but are not variables when they may be freely shifted about.

The question of whether you are plotting individual figures (no variable, one dimension), or series of figures (one variable, two dimensions), or series of series of figures (two variables and three dimensions), is therefore a matter of whether the arrangement of the component parts is fixed or not. If your individual figures (together with their stubs) can be shifted freely up or down in their places, you are only plotting several single figures, and need to use but one dimension. If however, their arrangement is fixed and dependent on each other, you are plotting one series of figures and need two dimensions. Likewise if your columns or rows of figures can be freely shifted about among each other, you are plotting several series and need only two dimensions, but if their arrangement is fixed and dependent upon each other, you are plotting a series of series and need three dimensions.

All of this may seem very confusing just at present, but as time goes by and you become more familiar with the nature of your figures, you will begin to understand the interrelation of these figures and then you will be able to use the rule just considered to determine quite arbitrarily in advance the type of chart you will need for illustrating them to the best advantage. We will therefore table this matter for the present, with the purpose of returning to it later on with a fuller understanding.

A corollary of this rule, however, you can keep and make full use of from the outset. That is, never to use more dimensions than are necessary for your chart. Do not use two dimensions when one will serve your purpose, nor three when two will do. If, through an excess of zeal, you violate this rule, and use more materials than are necessary, you will merely defeat your purpose and confuse the reader of your chart.

Let us take a simple example. Suppose we wish to illustrate with a chart the relative sizes of two cities, with populations 5,000 and 10,000 respectively. Obviously the second

town is twice as large as the first. There is but one variable present (none in the data-stubs or column-headings), the sizes of the cities. According to our rule we will use one-dimension charts for these figures, but as there are two of them, we will have to use two charts.

Because the charts are single-dimension or single axis ones, it is obvious that they will be straight-line ones. Two lines or bars, the one of which is twice as long as the other, will illus-

Fig. 58.

trate the two populations, so that the reader cannot help arriving at the conclusion that one city is twice as large as the other, merely from a glance at the chart. This chart will be both clear and accurate. It is therefore correct.

But suppose we had in this case used two-dimension charts. Suppose for example that we had used square areas instead of lines. And because the second city is twice as large as the first, suppose we therefore made each dimension in the second

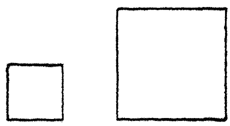


Fig. 59.

chart twice as great as the corresponding dimension in the first chart. You will already be objecting that the area of the second chart will not be twice as great as the area of the first, but four times as great. Look at the chart. Notice how surprising is the difference between the two cities as here represented.

And notice also that it is very hard, from the chart alone, to decide exactly what the ratio between the two areas is. The eye does not readily compare areas with precision. At first glance one might be inclined to say that the second area is five times as great as the first; though we who made it know that it is only four times as great. This illustrates a very important rule—that it is difficult to compare areas. That rule follows from the original principle, not to use more dimensions than necessary, but it is worth keeping in mind as a particularly important phase of the principle.

But you will say, "Cannot we show the areas in true proportion?" We can, if we will but take the trouble. A little figuring will show you that if the areas are to be in the ratio of 5,000:10,000 or 1:2, the sides of the squares must be in the ratio of one to the square root of two. Get out your pencil and paper—or your slide rule—and figure out the square root. It happens to be 1.414. Consequently, if the square areas are to be in the proportion of one to two, the sides of the squares

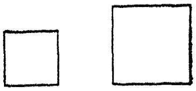


Fig. 60.

must be in the proportion of 1 to about 1.4. Now draw the two square charts in such a way that the sides of the second are 1.4 times as long as the sides of the first.

A glance at the resulting chart will disappoint you. Instead of making the difference in size clearer, you have apparently minimized it. The two areas no longer seem so very different in size. It will take a clever reader indeed to realize from a study of the two areas that the second city is fully twice as great as the first city. And the worst of it is that many readers will, through ignorance, fall back on the length of the sides as a basis of judgment, and decide that one city is only half as large again as the other. In short whether you use the sides or the areas as your own basis, there is a good chance that your reader will happen to use the other basis and so be entirely misled by your chart.

Obviously this is no less true when we use circles or other regular areas in the place of squares. A circular, like a square,

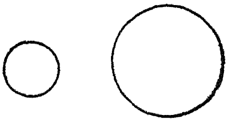


Fig. 61.

area varies with the square of its linear measurements. If you make the radius of one circle twice as great as the radius of the other, the first area will be four times as great as the first. If you make the areas proportionate, the radii must be in the relation of 1 to the square root of 2. Both circle and square require the more or less tedious computation of square

roots and repay this labor with inaccurate and ambiguous results.

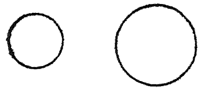


Fig. 62.

We can carry this example further, into three-dimensions. That is to say, suppose we attempt to show a one-dimension fact by a three-dimension chart. Suppose some bright young illustrator suggests that, since it is population we are showing, we use the picture of a human being for a chart. Now let us make the height of one person just twice the height of the other. What has happened to the volume or weight of the

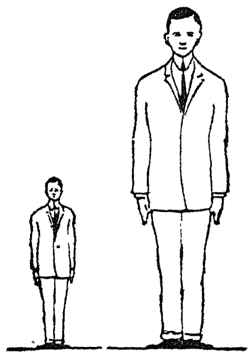


Fig. 63.

two persons—assuming that they are similarly shaped? Clearly one is eight times as large or as heavy as the other, for the volume of cubes or solid bodies varies with the cube of linear dimensions. Our pictures give a grossly exaggerated impression of the comparison of the two cities.

To counteract this, we can make the volume or weights of the two bodies represented in the proportion of one to two.

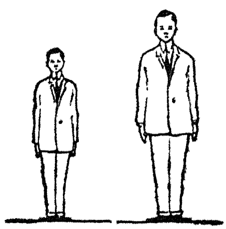


Fig. 64.

But in this case their heights and other linear dimensions are going to be in the ratio of one to the cube root of two. You will find if you look it up, that the cube root of two is 1.26. And you will be even more disappointed with your resulting volumes when you have drawn their heights to scale, than you were with the resulting surface areas. Meanwhile your poor reader will be trying to choose between three ways of judging from your picture—height, surface area, and cubic volume—with a two to one chance of making the wrong choice.

The illustration of population by areas or solids is a frequent type of faulty chart. We have all seen pictures purporting to show the sizes of various armies, for instance, by pictures of soldiers drawn to different sizes. Even the United States Census,[2] has turned out an entire volume of several hundred pages, containing many circles illustrating relative sizes. It is perhaps the commonest form of deceptive or ambiguous chart used—most frequently occürring when the author has tried to combine a picture of his items with a chart of their mathematical ratios.

When you see such a chart in the future, learn to ask yourself what part of this chart shows the true ratio, height, area, or indicated volume. In nine out of ten cases you will find, by looking at the data, that the height or linear dimensions are in true proportion. You will then realize that the area or indicated volume is grossly deceptive and you will not be

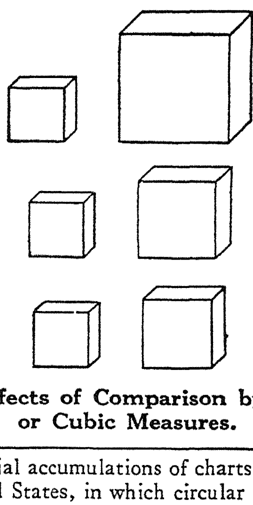


Fig. 65. The Effects of Comparison by Linear, Square, or Cubic Measures.

[2] One of the greatest official accumulations of charts is the *Statistical Atlas* of the Twelfth Census of the United States, in which circular areas were used very extensively with correct interpretation of values by areas.

misled by it. In any case you will quickly see that the man who made the chart either did not know his business or if he did, was guilty of shameless intent to deceive.

In short, the rule that no more dimensions or axes should be used in the chart than the data calls for, is fundamental. Violate this rule and you bring down upon your head a host of penalties. In the first place, you complicate your computing processes, or else achieve a grossly deceptive chart. If your chart becomes deceptive, it has defeated its purpose, which was to represent accurately. Unless, of course, you intended to deceive, in which case we are through with you and leave you to Mark Twain's mercies. If you make your chart accurate, at the cost of considerable square or cube root calculating, you still have no hope, for the chart is not clear; your reader is more than likely to misunderstand it. Confusion, inaccuracy and deception always lie in wait for you down the path departing from the principle we have discussed—and one of them is sure to catch you.

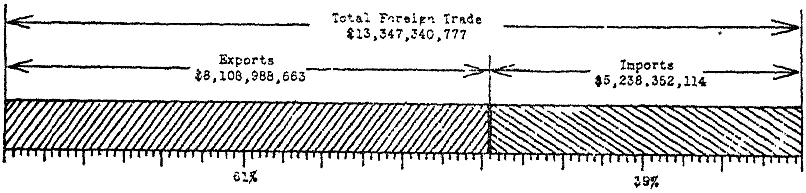
CHAPTER IX

HUNDRED-PER-CENT BARS

The division of a "whole" into its "parts" is logically one of the first steps in any analysis. Usually the graph illustrating this division belongs at the beginning of a statistical report. Thus, if your report covers the sales of the company, your first chart would break up total sales into the individual sales for each line or for each district. The remainder of the report, treating of details of the various "parts" (e.g., lines or districts) will then follow a summary chart which has established their relative importance.

A quantity can always be illustrated by a straight line, or, as it is commonly called, a "bar." Bars are the simplest and often the best form of graphs. The total length of the line then represents the total value of the quantity. When we speak of a line in charting, we do not mean an imaginary straight line having neither width nor depth, for that would be invisible and could not, of course, be actually used in illustrations. In its place we use the bar, with a visible width (and the actual depth or thickness of a layer of ink). But it is still proper to speak of this bar as being a line or one-dimension chart, for its width and thickness are constants, necessary to give visibility to the line, and its length alone is significant.

Now a single bar, illustrating a single figure, will have no particular meaning for the reader, because he has nothing to



FOREIGN TRADE OF THE UNITED STATES
1920

Fig. 66. A Simple 100% Bar.

compare it with. There is therefore but one case in which we have any use for the single bar, shown by itself. This is the case in which we wish to show the parts into which a total may be broken up, or of which it may be composed. And whatever the quantity or total be which is thus composed of two or more parts, we may call it 100 per cent of itself and the parts will then be certain percentages thereof. Hence, this single bar shown by itself, has come to be called the "100% bar," regardless of whether the total and part quantities be quoted in the data and chart in actual values or in relative percentages. Irrespective of scale or calibrations, the bar represents a whole or 100 per cent, and its divisions or parts represent parts or percentages thereof.

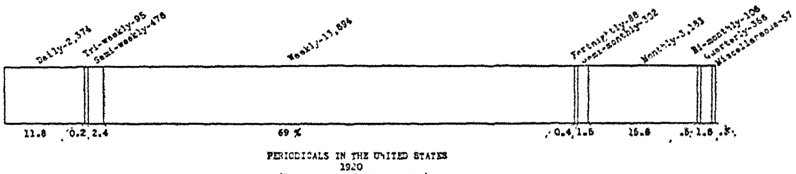
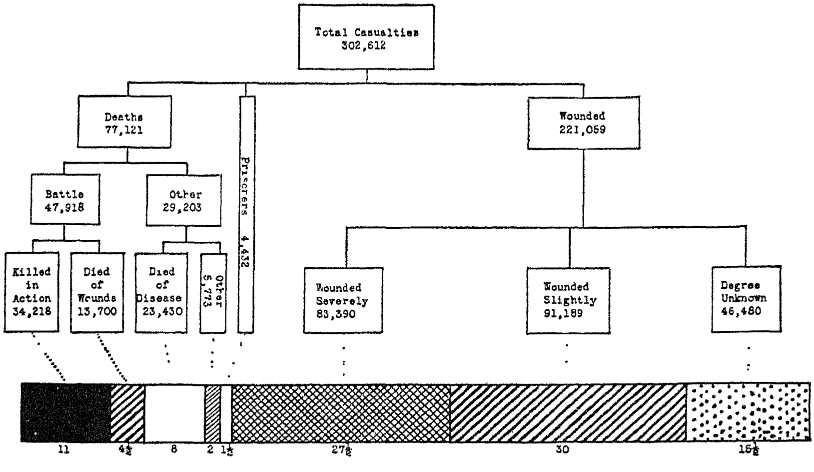


Fig. 67. Many Segments, No Shading.

In the 100% bar we have the mathematical equivalent of the classification chart. Turn back to Chapter III if you have forgotten what a classification chart is—or better still, don't turn back, but stop and try to remember it. The classification chart sets forth the ideological or schematic relation of things. Commonly it displays the parts of which a whole is composed. When it does this, it is very similar to the 100% bar. The difference between the two lies in the fact that the classification chart shows what the parts are but does not tell their relative importance or size, while the 100% bar shows their relative importance or size, upon the basis of certain numerical data.

Thus the best labelling (for labelling is half the work of chart-making) for the 100% bar would seem to be a classification chart placed, obviously, above the bar. As a matter of fact, wherever it is typographically possible, this is correct. Use a classification chart to label a 100% bar, or a 100% bar to illustrate a classification chart, and you have the best possible combination. Of course where some parts are relatively very small, typographical difficulties arise. If you cannot overcome these by setting certain words on edge, you will be



AMERICAN CASUALTIES IN THE WORLD WAR
1917-1918

Fig. 68. Classification-chart and 100% Bar.

obliged to content yourself with a general grouping of minor parts into a single part labelled "Other," "Minor," "Miscellaneous" or some such rag-bag title for stray odds and ends.

A further detail of the 100% bar and its labelling, is the scale. This should generally be in hundredths or percents. The data may be entirely in absolute quantities, but nevertheless the scale should show percentages. To prevent the confusion of scale and divisions of the bar, the scale should be outside the bar, and the best practice seems to be to indicate the scale by little notches or short perpendicular lines dropped below the bar, from its lower edge. The scale should have ten, twenty, or a hundred of these little lines, each indicating a division of ten, five, or one per cent. The purpose is to enable the reader, by counting notches on the scale, to compare parts of the bar with greater accuracy. For the same reason, the actual percentage to which each part is equivalent should be written or printed below the bar under the center of each part.

The bar itself, as has been said, should be of appreciable thickness. Too light or narrow a bar, such as a thin line, has no emphasis or force. But too wide or heavy a bar introduces two-dimensional rather than linear conceptions in the mind of the reader and sometimes produces undesirable optical illusions. The width of a bar should be such as to make it clearly visible at the distance from which it is to be viewed. The best

form of bar is generally between a tenth and a twentieth as thick as it is long.

The bar should be hollow, that is, outlined. The segments or parts of the bar may also be hollow, but it is better to shade them with distinct colors or shadings. Small dots, various hatchings (cross-lines), and double-hatchings (criss-cross lines), can be used to distinguish the various parts without using colors. But where colors can be used, they are sometimes de-

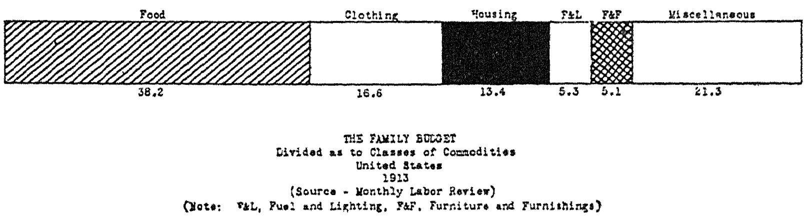


Fig. 69. Distinct Shading.

sirable, for the reason that solid shades are more forceful than black and white shadings. Care must be taken in either case, however, to see that the various parts are similarly emphasized by the color or shade, for otherwise one part will appear more important or larger than it really is. A solid black area will appear to be larger than a solid white one outlined in black, though really of exactly the same size, for the black is in itself so much more powerful than the white, and has further gained by absorbing its black outlines. Experiment will soon show

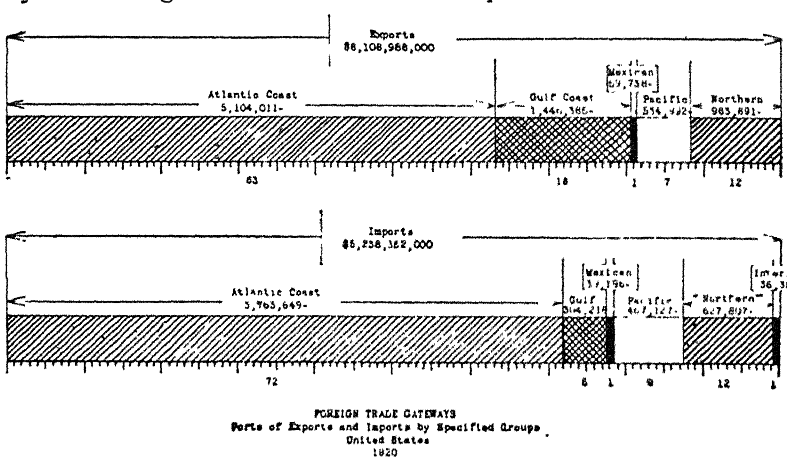


Fig. 70. Two Bars are Easily Compared.

whether an optical illusion[1] is being introduced by the shadings, and those combinations which will bring out the various parts equally.

The data for the 100% bar need be no more than a list of the parts of which the whole is composed, with their respective percentages, and either with or without their respective absolute quantities, according to your wish. If the quantities are important, or you think that someone is likely to call for them, add them and forestall criticisms; if they are unimportant, they can be safely omitted. While the percentages are almost always desirable and are best placed below the bar, as part of the scale, the absolute figures or data, if inserted, are best placed immediately over the bar, as part of the classification chart which is used for the labelling.

That data of this character calls only for a single dimension

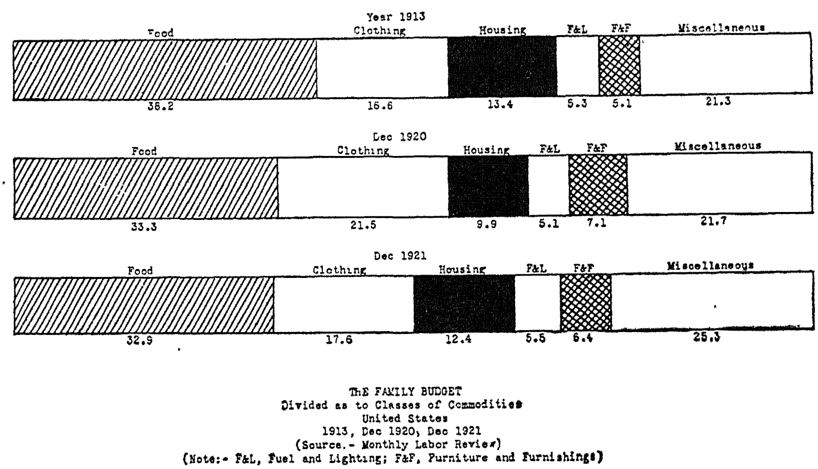


Fig. 71. Comparison of Three Different Years.

chart, according to the rule of chart dimensions is obvious. For the individual figures, with their stubs, in the table of data, can be shifted freely up or down the column or across in the row in which they are tabulated and the stubs therefore cannot be said to form a variable. In another chapter we shall consider another way of presenting the same data, in which each figure forms a separate bar, and the series of figures in

[1] For a discussion of the various optical illusions to be avoided, see Willard C. Brinton, *Graphic Methods for Presenting Facts*, Engineering Magazine Co., pp. 358, 359. See also Appendix D.

the data are shown by a series of separate bars. In the 100% bar you may, if you wish, think of these same bars as being again shown separately, only placed end to end instead of one above the other. And data of the 100% type, that is, in which the figures can be added together to form a coherent and significant total, is the only case of data which can be shown on the 100% bar. For all other series of figures the bar charts discussed in a later chapter must be used.

Chapter X

PIE-CHARTS

Throughout your study of charts you will find some which are more useful for popular consumption than others, but you will not find many which are more purely popular in appeal than the 100% circle or pie diagram. For analytical purposes it has nothing to recommend it, but for sensational values it is in general without an equal. If you are research-bent, you may safely pass by this chapter on popular exegesis, but if your object is advertising, you will seize it to your heart.

We have just seen how a single bar can be taken to represent 100% and can be cut up into segments or parts the lengths of which correspond to the relative sizes or percentages of the various parts of the 100%. The fact that the line is a unit, and so long as it remains the whole, can never be more than a unit or 100%, should suggest something. It should suggest that the total length of the whole line is relatively unimportant. It is unimportant because the reader is not asked to compare the total length of the line with the total length of any other line. There is no other line to compare it with, unless a second 100% bar is lying around handy, in which case the second would presumably have the same length, because it too represents 100%. Hence, you will say, why have any total length at all?

Centuries ago it was a moot question among philosophers, whether the Lord could make a yardstick which was endless. Then someone suggested that the yardstick be bent into circular form and the question was dropped. Let us perform the same operation on a 100% bar. Imagine, if you wish, that the bar is so very thick for its length that while one edge becomes the circumference of the circle, the other shortens down to and becomes the center of the circle. Division lines between the component parts of the bar become rays or radii of the circle and serve to mark off the corresponding component seg-

ments of the area of the circle. Here you have in a nut-shell the pedigree of the pie-chart.

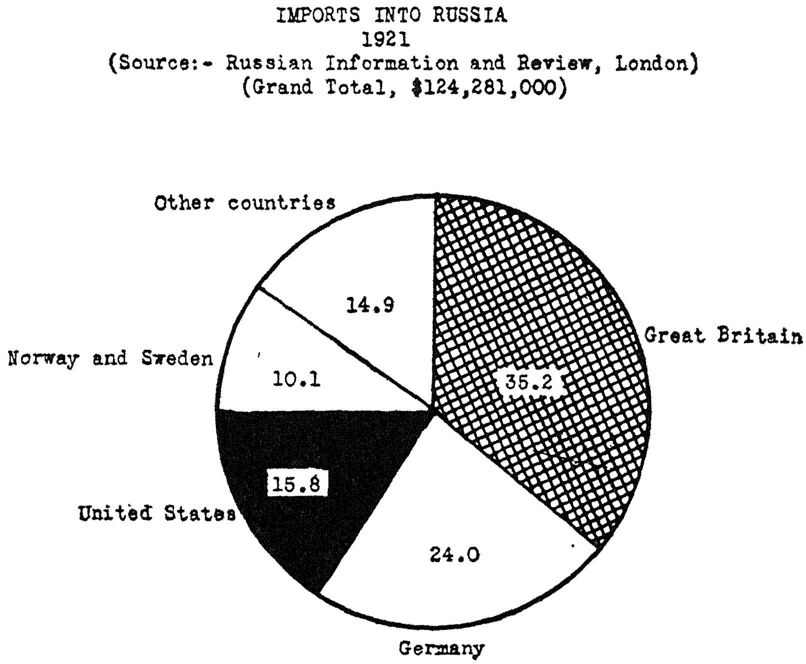


Fig. 72. A Simple 100% Circle.

It is now time to let you into the secret that the rule of dimensions of charts, which you doubtless memorized in a previous chapter, has apparent exceptions. The pie-chart is one of them. For few readers will judge quantities by either the arc at the perimeter of the circle or the subtended angles at its center—on the contrary most of them will judge entirely by the areas of the segments. In short, the pie-chart appears to be a two-dimension (area) chart used for one-dimension data. The fact is, however, that, as in the case of the 100% bar, the area of the chart varies directly with one dimension, the other dimension being constant. In the 100% bar the width of the bar was constant in the 100% circle the radius must be constant for all circles compared. Then the area of the segments varies directly with their arcs or angles and the chart has but one significant dimension. It is only an apparent exception to the rule.

The disadvantages of the pie-chart are many. It is worthless for study and research purposes. In the first place, the human eye cannot easily compare as to length the various arcs about the circle, lying as they do in different directions. In the second place, the human eye is not naturally skilled at comparing angles—those angles at the center of the circle, formed by the various rays or radii and subtending the various arcs. In the third place, the human eye is not an expert judge of comparative sizes of areas, especially those as irregular as the segments of parts of the circle. There is no way by which the parts of this round unit can be compared so accurately and quickly as the parts of a straight line or bar. Moreover, when, as frequently happens, several pie-charts are shown together, the various slices in one chart cannot be so easily compared

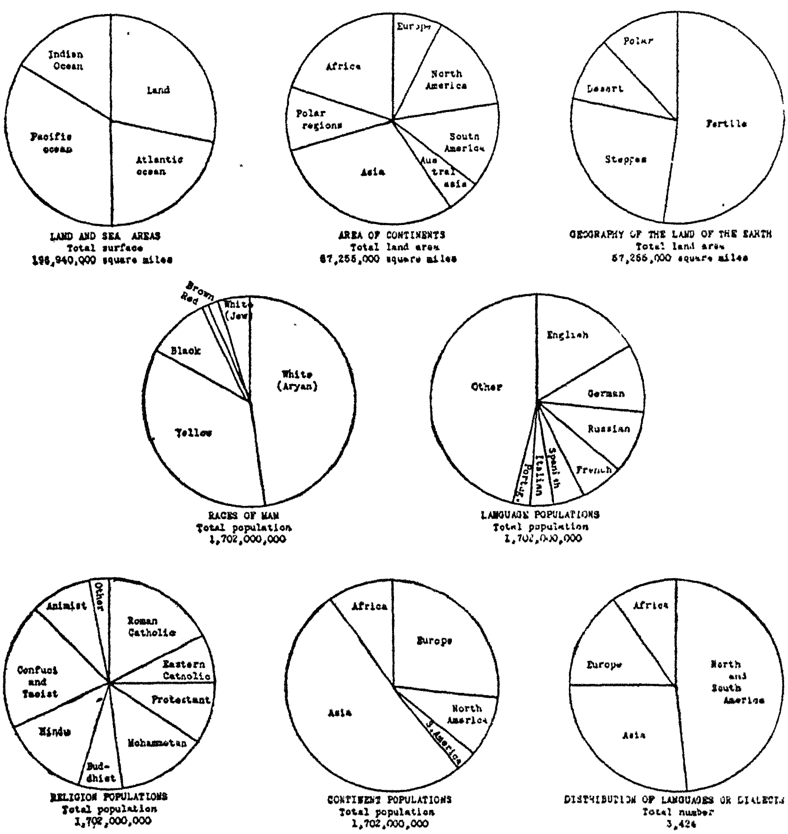
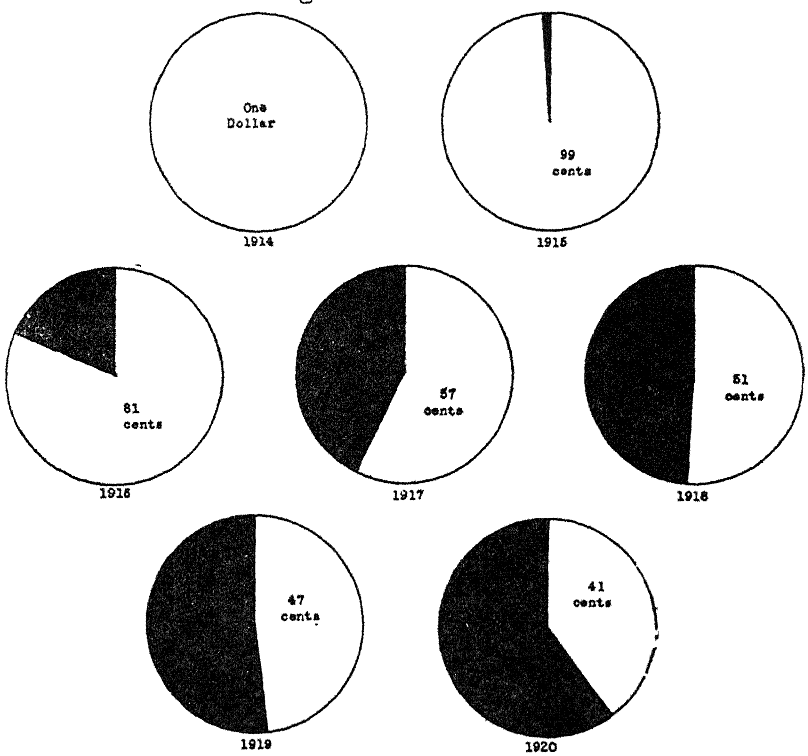


Fig. 73. Accurate Comparisons Cannot be Made.

with the corresponding slices in the next, as can the various parts of one 100% bar with corresponding parts of another bar. The two bars can be placed one above the other, so that comparison from one to the other can be made at once, but no arrangement of the two circles will make comparison so simple.

In the labelling of the pie-chart, you will furthermore encounter typographical difficulties. It is not ordinarily a good thing to make a reader crane his neck at various angles to read writing along every point of the compass, so you should not, as so many do, write on radii from the center of the circle. On the other hand, unless the chart and its segments are very large as compared with the size of the printing, you will introduce tricky optical illusions if you write all labels in the same directions inside the segments.



PURCHASING POWER OF THE DOLLAR
OF 1913
when used for food at retail
U. S.

Fig. 74. **The Less Detail, the Better.**

Sometimes, the best rule is to put the labelling away in a key or explanatory list of the shades or colors of the various segments. Only if the labels are very brief, and your segments are all large, can you stow the labels into the segments without greatly altering their apparent sizes. When neither plan is feasible, and you feel that you must have each segment, however small, immediately labelled, place the labels outside the circle, adjacent to the proper segments, with the printing in as nearly the same general direction for all as you can arrange, so that the apparent sizes of the segments are not confused by printing and the reader need not climb around the edge of your chart to decipher it.

As a general thing, however, there is one part of the labelling which can always be attached to the chart, namely the figures of the percentages for each segment, which in the bar-

RETAIL FOOD ESTABLISHMENTS
NEW YORK CITY
1921
(TOTAL NUMBER, 75,412)

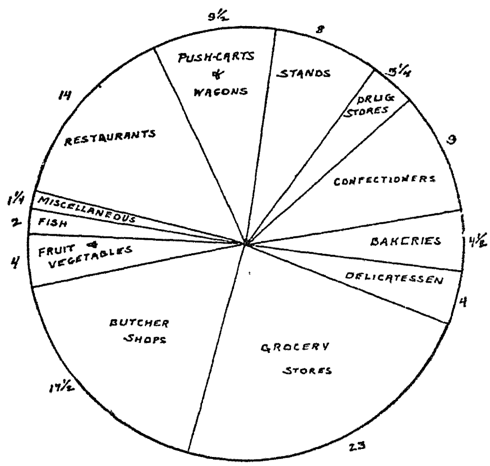


Fig. 75. Labelling is Difficult.

chart were placed immediately below the bar. These figures should always be placed close to the segments, but usually outside the circle, so that the reader who wishes to have the precise percentage figure represented by a segment, can always do so.

The scale (without scale-figures) may be placed inside the circle and unlike the 100% bar may or may not show in the finished chart. Special paper is marketed for these charts with

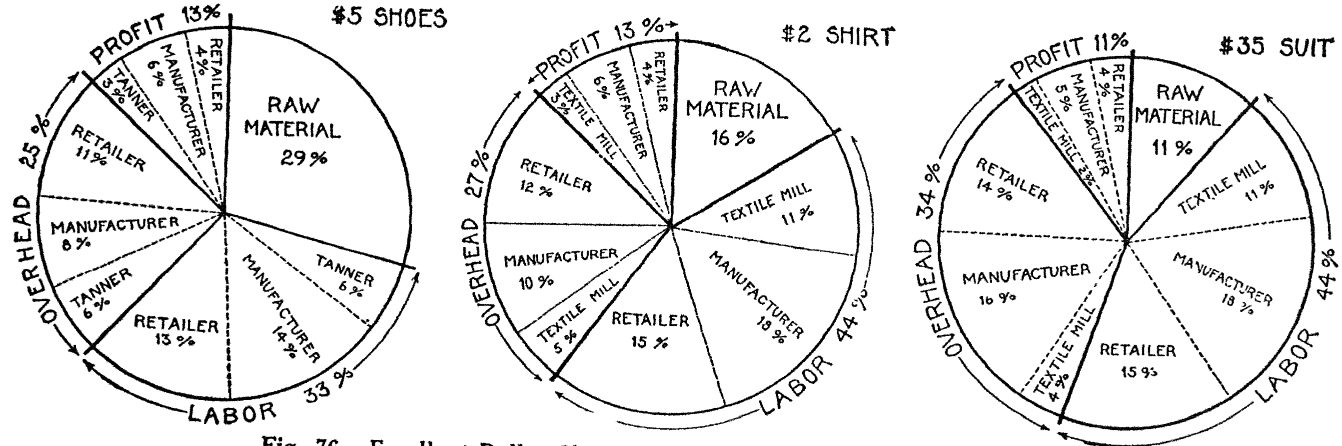


Fig. 76. Excellent Dollar Charts.

Analysis of the cost to the consumer of a $5 pair of shoes, a $2 shirt, and a $35 suit of clothes, carried through three processes: preparing leather or cloth from the raw leather, cotton or wool; manufacture of the article of clothing; and selling it at retail. Direct sale from manufacturer to retailer is assumed.—*Permission of Mr. Carl Snyder.*

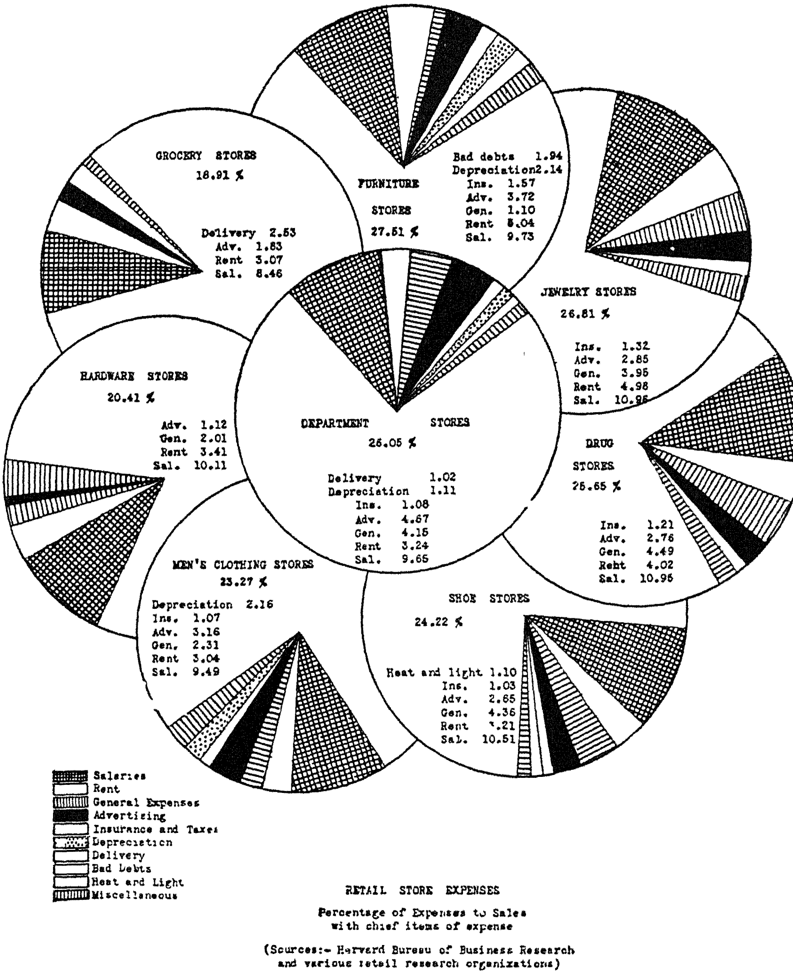


Fig. 77. Shading the Segments to Increase Popularity.

the circle printed in and already divided into a hundred parts by small notches within the circumference. The use of this paper will save you much time if you wish to make the segments accurate in size. It is tiresome to use a protractor marked off into 360 degrees, and to calculate the decimal equivalent or percentage in degrees. Unfortunately, the metric system has not been applied to circular measurement so as to give any circular or angular measurement which employs a decimal system.

The advantage of the pie-chart is psychological. It instantly commands the reader's attention. A circle is, of all geometrical patterns, the easiest resting spot for the eye. The fact is well known to advertisers, who frequently use circles and circular outlines to draw attention to their advertisements. Hence if your chart is designed for publication, or for presenta-

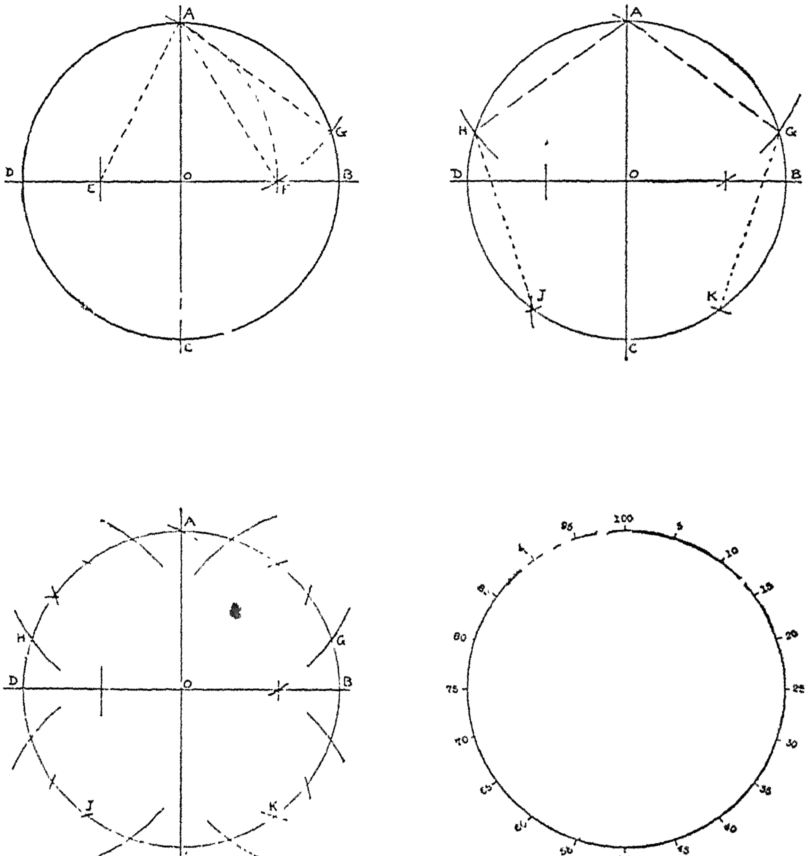


Fig. 78. How to Calibrate the Circle.

Bisect OD at E and project EF equal to EA. Project the chord AG equal to AF. With dividers set for this distance, lay off from A the points, G, H, K, and J (as in upper right-hand diagram). Similarly lay off four points from B, C, and D (as in lower left-hand diagram). Erase all other marks and calibrate the twenty points so found, to form the finished circle (as in lower right-hand diagram).

tion to readers whose attention may be easily diverted, you will find the pie-chart a powerful means for presenting your facts. Attention will be focused upon it at once, and it is as

simple to understand as its name—far too simple for anyone to misunderstand. Because it is circular, there is no question but that it represents a whole and the various slices of the pie belong to their respective items.

COST OF THE WORLD WAR TO THE UNITED STATES
ESTIMATE ON JULY, 1921
GRAND TOTAL COST—$50,168,625,707.16
(SOURCE: WORLD ALMANAC)

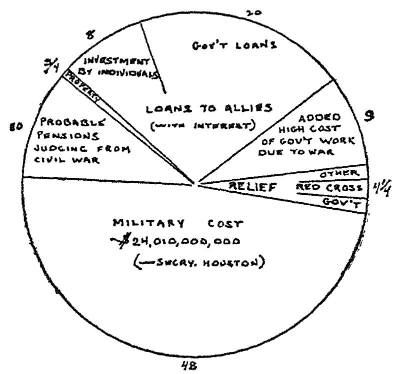


Fig. 79. Many Segments, No Shading.

A very sound use of the pie-chart occurs in the case of financial data. Here the whole circle or pie can be spoken of as a "dollar," and the various segments, the parts of the dollar, so many cents (or percentages) each. Charts of this type have been used to show the distribution of costs in a plant, or the parts of a financial budget, or the shrinkage of the dollar in high-cost-of-living studies. Through the fortunate coincidence that our metal currency is round and our dollar divided into a hundred cents, the shape and the labelling of the chart both find immediate understanding in the mind of the reader.

Fig. 80. A Pie-chart in Metal.

The "Swift Dollar," as it was called, on one side of which a chart shows the division of income from sales.

The pie-chart must therefore be accepted as an advertising medium of value. It has strong popular elements. But it has

no place in the statistical workshop, or research laboratory. Before using it in the place of the simpler and sounder 100% bar, you should carefully gauge your audience or readers, and only if you believe that you have begun to strain their interest should you judiciously insert the pie-chart. In a sense, it might be construed as an insult to a man's intelligence to show him a pie-chart, but the insult is not often resented. For if your main object be to get a story across, you are justified in taking that means which will encounter least resistance, and in making your story as simple as possible. For publicity purposes, the pie-chart is therefore almost invariably better than the bar.

CHAPTER XI

BAR-CHARTS

A series of quantities or values can be most simply and often best shown by a series of corresponding lines or bars. All bars being drawn against one and the same scale, their lengths vary with the amounts which they represent. In a previous chapter, the 100% bar was described, in which a single bar, whose total length had no significance, was divided into parts in order that the relative size of the parts might be seen. In this chapter we propose to use several bars, which are not divided up into parts, but which can be compared as to their total lengths, in order that their relative sizes may be seen. While this new method could be employed with the same data, it is generally more useful for data in which the various items are not being shown as parts of a total, but as individual and co-equal totals in themselves.

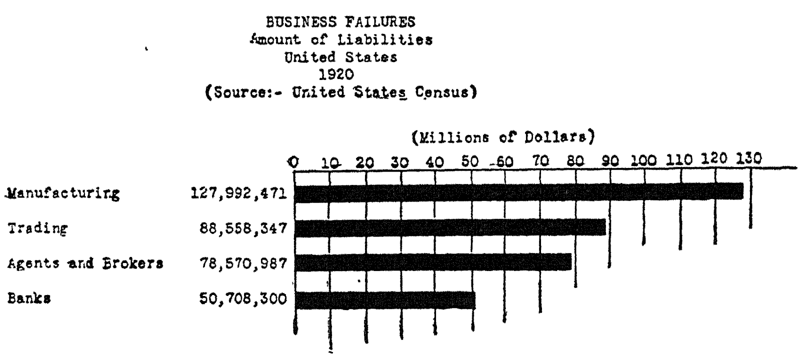
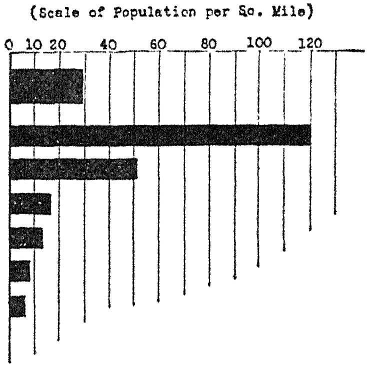


Fig. 81. A Simple Bar-chart.

Bar-charts are most flexible and can be varied to suit the individual whims of the maker. In general, however, there is one style or form which will be found most satisfactory. It consists of a horizontal grouping of bars alongside of the data. The chart is arranged in tabular form, with items or stubs in

a column to the left, with figures in a column beside the stubs and with bars in a column beside the figures. Several columns of figures are sometimes desirable, just as in the table of data, to show sources or original figures from which the charted

	Population	Area square miles	Pop. per sq.mi.
TOTAL	1,702,520,000	57,255,000	29.6
Europe	464,681,000	3,873,000	120.0
Asia	872,522,000	17,206,000	50.7
North America	150,000,000	8,589,000	16.3
Africa	142,750,000	11,623,000	12.3
South America	56,340,000	7,570,000	7.4
Australasia	16,230,000	3,313,000	4.9
Polar Regions	...	5,082,000	...



DENSITY OF POPULATION OF THE EARTH
by continents

Fig. 82. Detailed Data may be Included.

figures are obtained. In any case, the bars should represent the most important set or column of figures, and there should be normally but one column of bars.

There should be but one column of bars because the bars can be advantageously compared only when they are side by side, one below the other. They should all begin at a uniform point or distance from the figures, so that their lengths can be compared out at the far ends. They should be the last column on the page, because their uneven lengths make further columns wasteful of space, and the addition of further columns introduces optical illusions which should be avoided. If a column of bars were to be followed by a column of figures, the reader would be apt not to compare the lengths of the bars, but their shortage from the last column.

It is a very common error to place the data inside the bars, for by so doing the reader is led to compare those parts of the bars which are clear of figures, rather than the entire lengths of the bars. This optical illusion exaggerates the difference in lengths of bars. Another mistake which is often made is to place the data out at the varying ends of the bars, for here the reader is led to compare the lengths of bars plus data, rather than of bars alone. Here the optical illusion minimizes the

difference between bars. The proper place for data is in a column at the left of and immediately before the bars themselves, with no more reading matter to the right of the bars.

The scale for a bar-chart should be placed at the top of the chart, immediately above the uppermost bar. A field in fainter color (green is most useful for chart-fields) or thin lines should be drawn into the chart by extending down from the scale a few lines which mark off convenient distances on the scale. Thus the reader is enabled to compare lengths of bars far distant, by noting their relative positions against the field or background. Care should be taken that the lines of the field do not cross the bars, else the field will cease to be a background and will become a screen in the foreground.

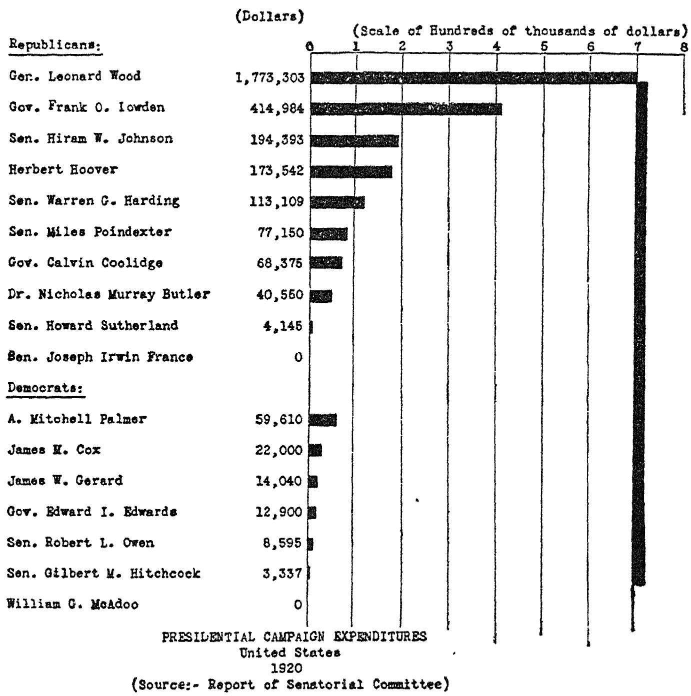


Fig. 83. A Long Bar Broken to Save Space.

The bars should be of uniform thickness or width, as it is the variation in their length which is significant and variations in width would produce area-illusions. This rule is obvious—so obvious that where a large number of bars are shown and the reader is already thinking in terms only of the lengths of

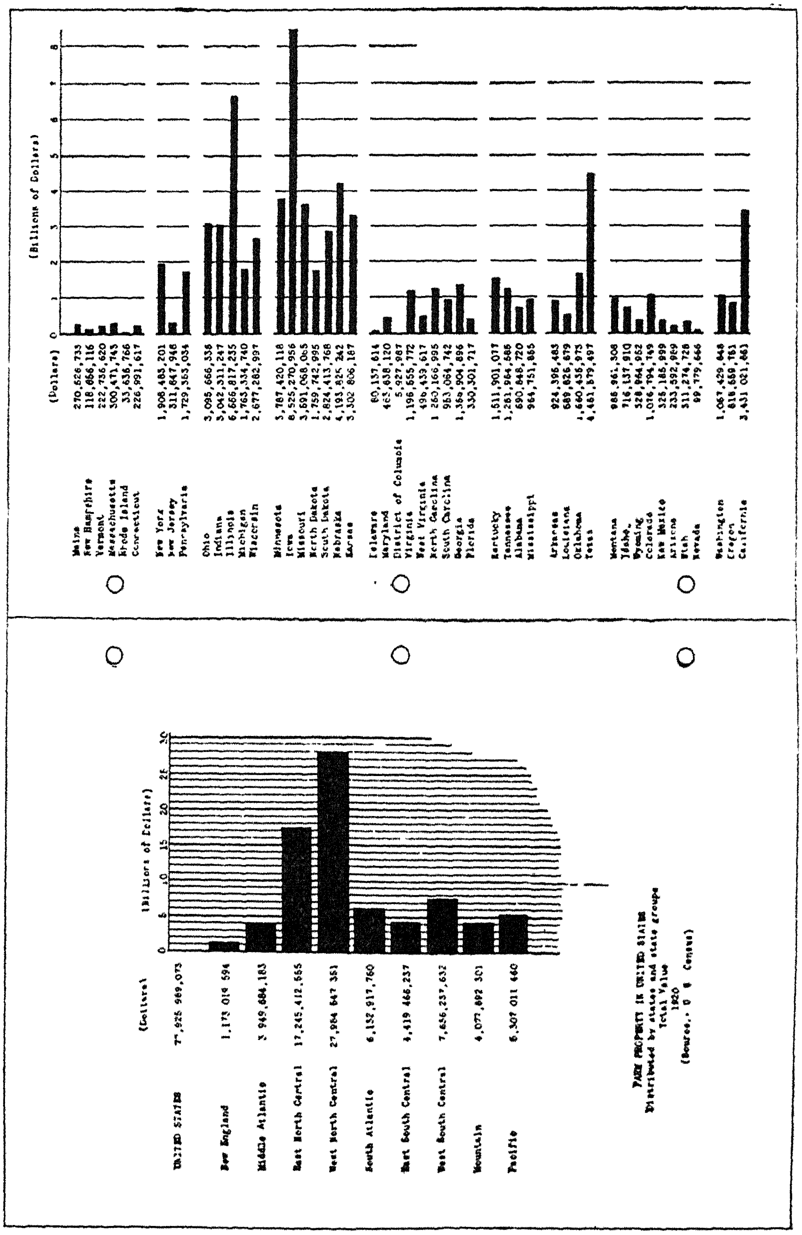


Fig. 84. National Distribution by States and State-groups.

the bars, it is permissible to violate the rule to emphasize important or group-total bars. The simplest example of the extra-wide bar is an average or total bar for the entire series.

The extra width given this bar need be only slight to make it stand out from the rest as a sort of type or normal against which the individual bars can be measured. And this should be done only when there is no danger of the reader's attaching importance to areas.

It sometimes happens that you desire to show two tables on the same subject, one giving a large number of individual items and the other a few sub-totals. It is a good plan in such cases to place the two on separate sheets that face each other, so that they both show at the same time, the one giving summaries of the other in the form of sub-totals. Where the summaries are not averages, but are true summations or

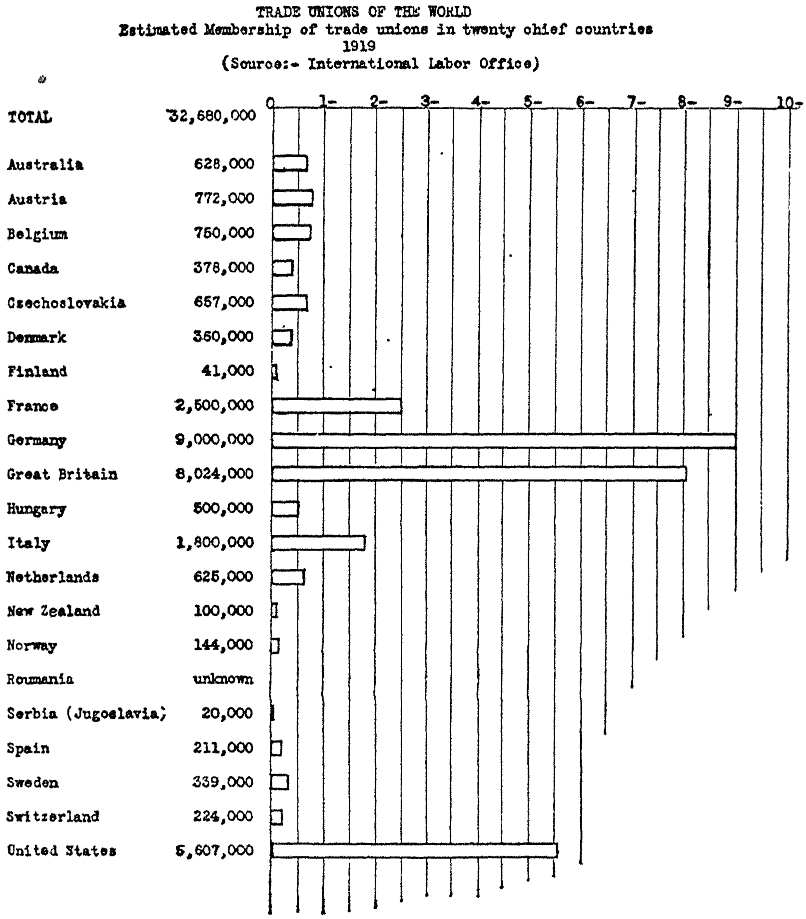


Fig. 85. An Alphabetic Arrangement.

totals of the items they include, you will find it well-nigh impossible to draw both charts to the same scale, without having one or the other so large or so small as to be useless. You will therefore be obliged to shorten your scale, that is, use a smaller scale for the summary-bars than for the individual bars.

In such cases it is not a bad plan to make the sub-total bars wider. In exactly the same proportions as you condense the scale, you should thicken the bars plotted thereon, so that the areas of the summary-bars will be equal to the combined areas of the individual bars of which it is the total. This is one case in which a slight use is made of the conception of area, or two dimensions, but it is negligible, for the reader is not called upon to compare areas—only lengths—in all cases, and the thicker bars remind him that the scale has been shortened on the sub-total chart and prevent his making confusing comparisons between the two charts. In short, when the group item in a series is an average of individual items in the series, it can be shown on the same scale, but where it is a total or sum of individual items, it can not be shown on the same scale without making the individual items small, but can be shown on a separate chart in reduced scale and with correspondingly increased thickness.

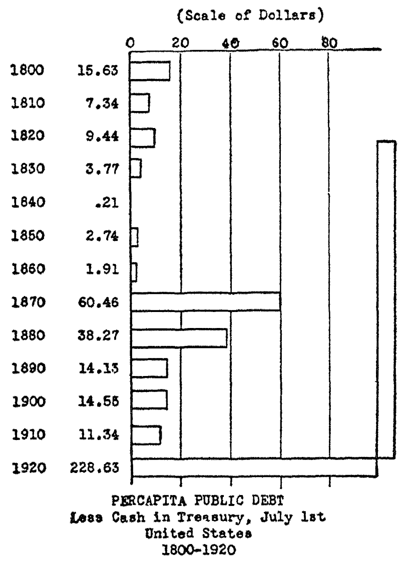


Fig. 86. Historical Data Must be in Order.

The arrangements of items in a bar-chart should be a matter for careful study, and no arrangement should be chosen which is not the best adapted for the special purposes of the chart. It is impossible to lay down absolute rules, for each case varies, and the author of the chart must rely on his own judgment rather than on rules of thumb. But in general the principles explained in a previous chapter on work sheets apply here, as the bar-chart is closely akin to the mere statistical table. When the items are historical, the earliest dates should be at the top and the last at the bottom of the chart,

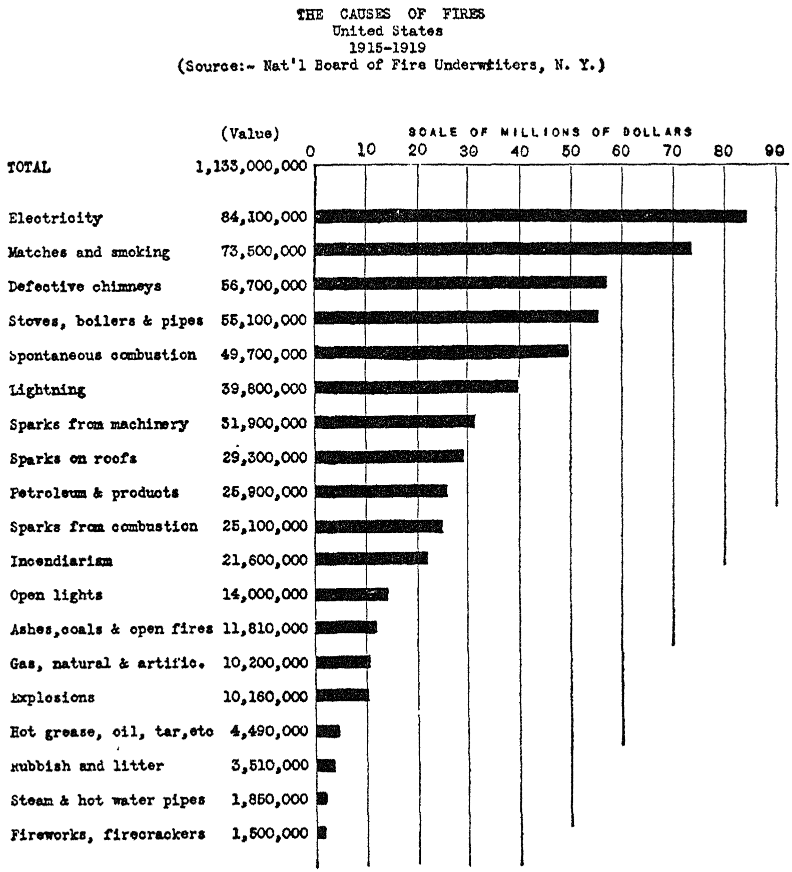


Fig. 87. Placing the Most Important First.

the order being strictly chronological. When the items are geographical, as for example a list of the States in the United

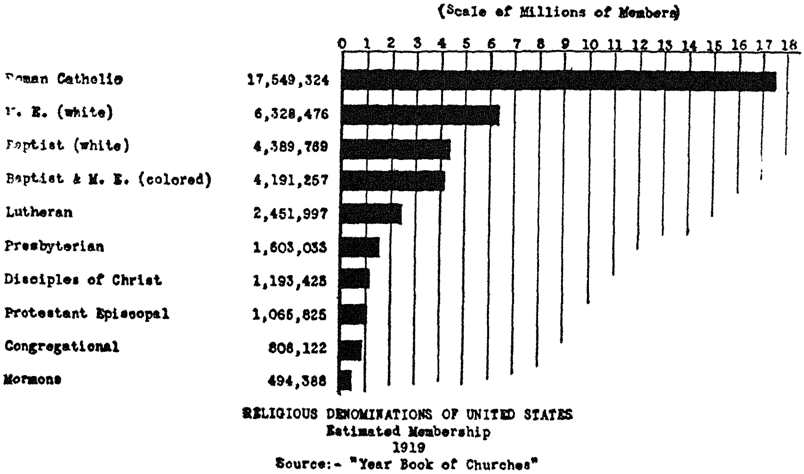


Fig. 88. The Arrangement in Order of Size is Popular.

States, a geographical arrangement (either in the order given in the Census volumes or in some specially designed order) is

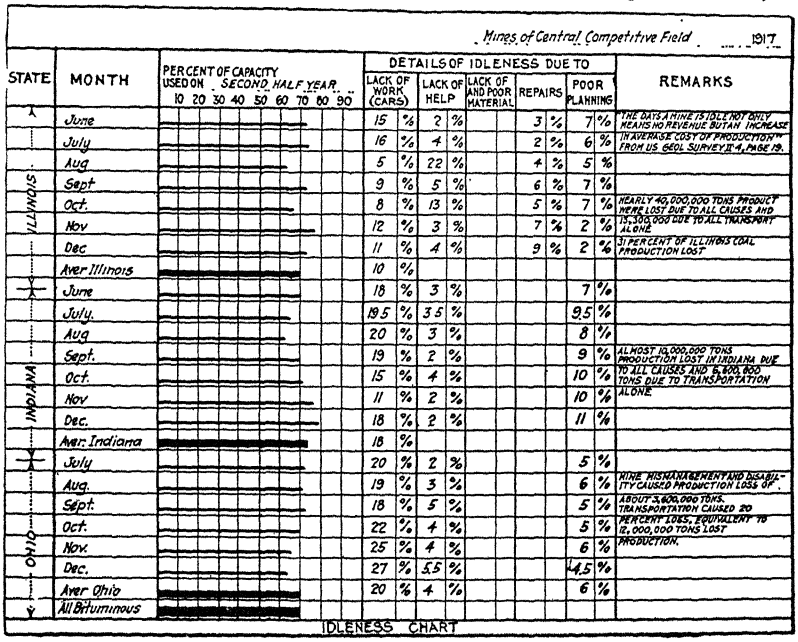


Fig. 89. The Gantt Idleness Chart.

Showing the failure to produce to capacity, together with data analysing the failure. This chart is one of a series of such charts on the coal mining industry, appearing in *The Dial.—Reproduced from "The Life and Work of Henry L. Gantt," American Society of Mechanical Engineers, paper of Mr. Polakov.*

preferable. A mere alphabetical arrangement of the States would have little to recommend it, as the reader of the chart must be presumed to have sufficient intelligence not to require a dictionary of the country. The popular arrangement is in the order of magnitude of the data presented, so that the longest bars are at the top, but unless the reader's sole purpose is to glimpse the names of the leading States, this arrangement is useless, as it lacks comparability with other charts. In general the arrangement of the items should be such as to afford the greatest aid to such analysis and comparative study as the chart may be subjected to.

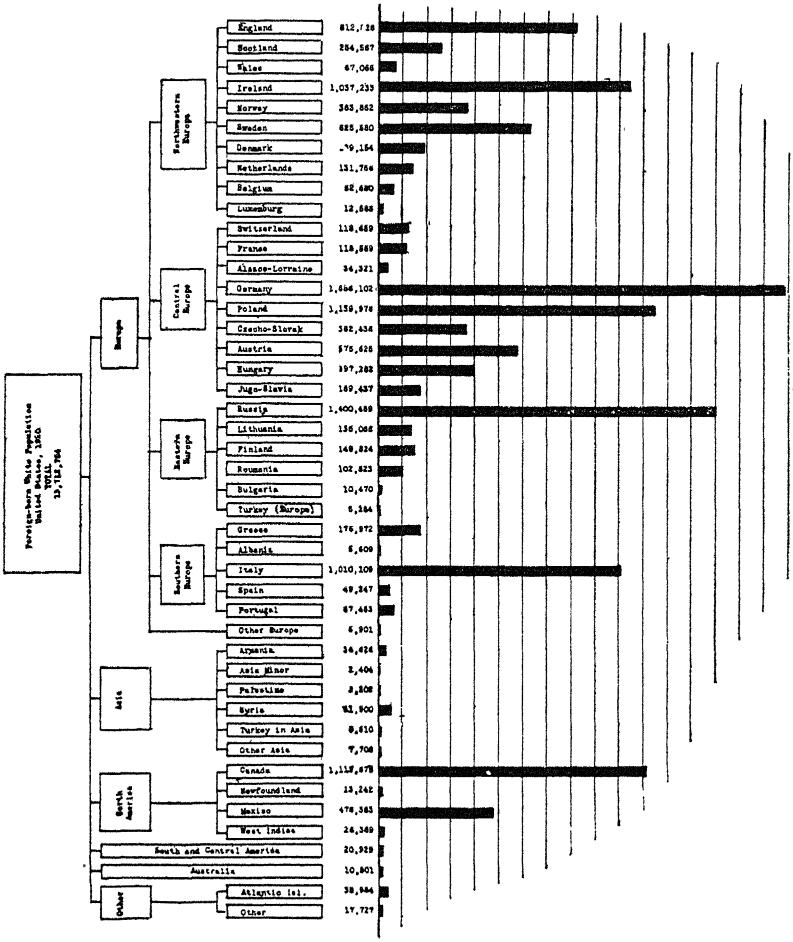


Fig. 90. Classification-chart and Bar-chart.

The bar-chart should be provided with column-headings for the data just as the work-sheet, but even more carefully prepared. Small keys or samples of the bar-chart can be in-

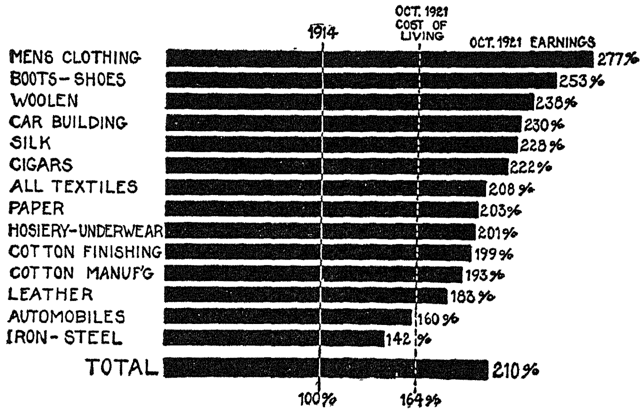


Fig. 91.

Average weekly earnings of workers in representative establishments in different industries throughout the United States, and the cost of living. (1914 = 100%).
—*Permission of Mr. Carl Snyder.*

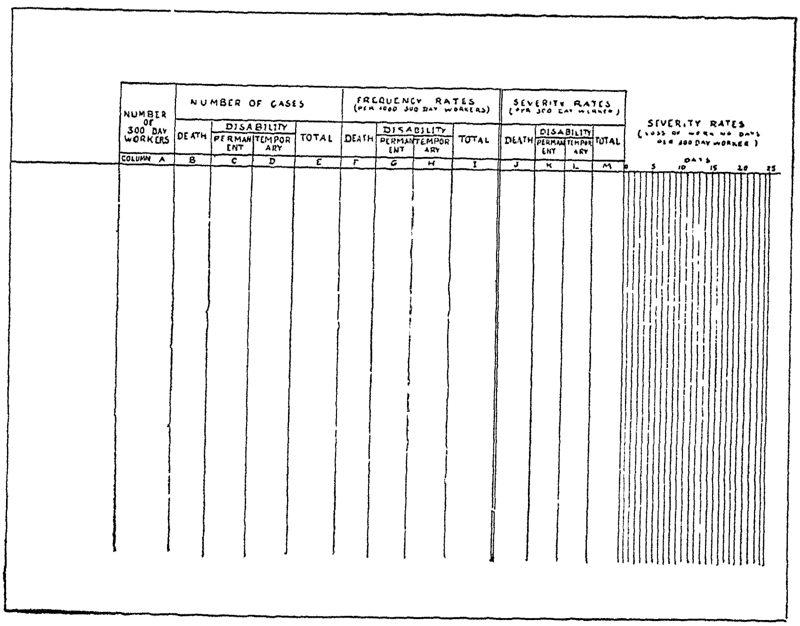


Fig. 92. An Office Record Form to Include Bars.

serted under the heading of the data which is plotted to show to which figures the bars belong. The scale of the bar-chart, above described, forms the heading for the bars themselves. When a series of charts are shown together in the same report, the charts of absolute figures can be displayed in typing and bars of one color, such as black, and the charts of relative percentages, per capitas, or averages can be distinguished by typing and bars of another color, such as red. In general a judicious use of colors will assist the column-headings and titles of the charts in differentiating charts in large sets.

The technique of bar-charts is so simple and they are so very effective, that they should be used freely in printed text-matter. No drawing or plates are needed. Printers have "rules" as they call them, which can be used to make solid bars, and these rules can easily be set up together with the type. The scale and field can be omitted and the bars alone will effectively tell the story of the main figures in the table. The combined table and chart can be used in printed text just as well as the table alone.

	Number of 300 Day Workers	Number of Cases				Frequency Rates (per 1000 300 Day Workers)				Severity Rates (per 300 Day Worker)			
		Death	Disability Permanent	Disability Temporary	Total	Death	Disability Permanent	Disability Temporary	Total	Death	Disability Permanent	Disability Temporary	Total
	Column A	B	C	D	E	F	G	H	I	J	K	L	M
Total	115,703	37	411	13 199	13,647	30	3 60	114 10	118 00	2 90	1 60	1 10	6 60
Ships	6,615	3	15	1,422	1,440	5	2.3	215 0	217 8	4 1	1 6	2 3	8.0
Mining	3,994		12	755	767		3 0	187 0	192 0		1 1	1 9	3 0
Cranes	4,362	1	15	813	829	.2	3 4	186 4	190.0	2 1	5	1 6	4 2
Steel	2,692	1	16	438	455	.4	5 9	162 7	169.0	3.3	2 9	1 5	7 7
Engines	31,249	22	160	4,348	4,530	7	5.1	139 2	145 0	6.3	2.9	1 4	10 6
Electrical	35,674	5	100	3,455	3,560	1	2.8	96.8	99 8	1 3	1 1	1 0	3 4
Power	2 226		9	186	195		4.0	85.6	89.6		2 1	1 1	3.2
Tools	24,359	3	68	1 486	1,557	1	2.8	61 0	63 9	1 1	8	6	2.4
Unclassified	4,552	2	16	236	314	.4	3.5	65.0	68 9	4 0	1 1	7	5.8

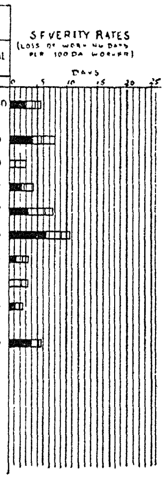


Fig. 93. Bars as Part of the Office Record.

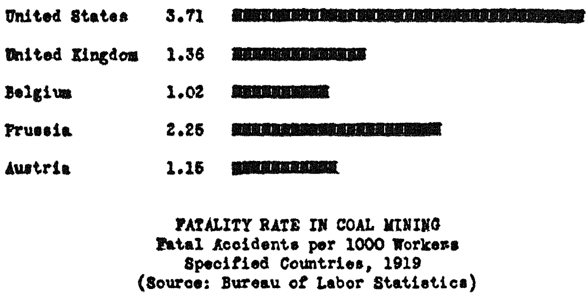


Fig. 94. Typewritten Bars for Typed MS.

When manuscript is typewritten, the bars can, if desired, be typed in, using such letters as "x" or better still, "x," "o" and "m" printed one over the other. This obviates the need of drawing, and illustrates excellently any tables which you are obliged to insert in the text. It has the further advantage of reproducing on carbon[1], hektograph, or mimeograph copies (though the latter can be secured also in drawings by the use of a special stylus).

[1]Not only can various combinations of colors in typewriter ribbons, be obtained, but also various colors in carbon paper, which can be successively inserted in the typewriting machine to reproduce the desired effect. Black and red are, however, usually sufficient.

CHAPTER XII

COMPOSITE BAR-CHARTS

What man has done once, he can do again, and since we have put several single bars together to make a bar-chart, we can put several bar-charts together to make a compound or multiple bar-chart. The single bar represented a single figure, the simple bar-chart a series or column of figures, and the compound or multiple bar-chart will illustrate a series of series (or columns) of figures. For the sake of convenience, we can divide these last into two classes, one of which we may call the compound and the other the multiple bar-chart. Nomenclature is of little importance but a precise use of names will help to distinguish two radically different forms.

Where each bar in a bar-chart is divided into parts, as was the single 100% bar, the name compound bar-chart is suggested. In such cases each bar is really a 100% bar by itself,

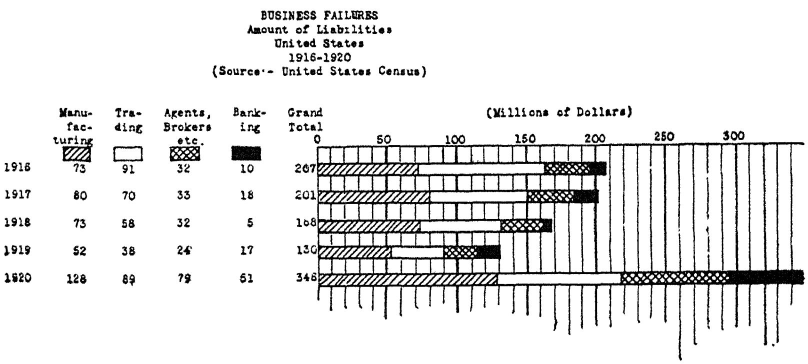


Fig. 95. The Compound Bar-chart.

but its length may be no longer constant and uniform, and may be made to vary in the fashion of the simple bar-chart. Its scale and labelling follow the same form. Data again should be at the left of the bars. The scale should be above the bars with a field projecting it downward behind them.

DIRECT COST OF THE WAR

National Debts of Chief Belligerents, in 1919

(Source: - E.M. Friedman, "International Finance")

	Pre-War Debt	War Debt		Total or Post-War Debt		
		Aggregate	Percent Increase	Aggregate @ Mint Parity	Per Capita @ Mint Parity	Per Capita @ 1921 Exchange
	$	$	%	$	$	$
TOTAL	28,669,000,000	183,272,000,000	631	211,941,000,000	329	
Germany	1,165,000,000	38,842,000,000	3,370	40,007,000,000	590	42,500
Great Britain	3,458,000,000	34,199,000,000	990	37,657,000,000	817	930
United States	1,208,000,000	25,389,000,000	2,164	26,597,000,000	249	249
France	6,598,000,000	23,896,000,000	423	30,494,000,000	768	1,640
Russia	5,092,000,000	20,314,000,000	393	25,406,000,000	138	INFINITE
Austria	2,631,000,000	14,440,000,000	548	17,071,000,000	551	745,000
Italy	3,031,000,000	11,978,000,000	392	15,009,000,000	408	1,525
Hungary	1,602,000,000	7,307,000,000	456	8,909,000,000	416	600,000
Australia	93,000,000	1,526,000,000	1,616	1,619,000,000	326	371
Turkey	667,000,000	1,335,000,000	201	2,002,000,000	94	
Canada	336,000,000	1,248,000,000	372	1,584,000,000	189	189
Belgium	722,000,000	1,167,000,000	162	1,889,000,000	247	568
Bulgaria	171,000,000	987,000,000	578	1,158,000,000	210	
Greece	188,000,000	333,000,000	177	521,000,000	105	450
New Zealand	446,000,000	288,000,000	65	734,000,000	632	720
Japan	1,261,000,000	23,000,000	2	1,284,000,000	2[illegible]	28

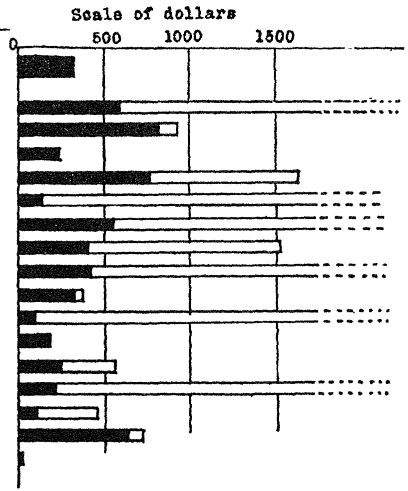


Fig. 96. The Chart Does Not Suffer from Detailed Statistics Attached.

The bars should be placed beside and on the same line with the data which they illustrate. As to the arrangement of the columns of data, it is best to have the column of totals, which the entire bar represents, at the beginning or end, and then beside it, in the order in which the parts will appear in the bars, the columns of various parts. Shading of the parts will follow the rules given in the chapter on 100% bars.

We will find two distinct types of the compound bar-chart, each belonging to a certain type of data. If the data expresses the values in absolute quantities, the totals for the various items or stubs (lines) will not necessarily be equal—in fact will very rarely be equal. So that the chart of this series will have bars of unequal total lengths. (Needless to say, the widths of bars will always be constant and uniform, save for the special exceptions noted in a later chapter.[1]) Such a chart has its scale and field measured off in units of absolute or

COAL RESERVES (Unmined) OF THE WORLD

1920 Estimates in millions of tons

	Total all kinds	Anthracite.
GRAND TOTAL	7,460,506	542,103
United States	3,538,506	16,163
Canada	1,361,000	2,000
China	1,097,000	427,000
Germany	467,000	-
Great Britain	209,000	12,000
Siberia	192,000	-
Australia	185,000	-
India	87,000	-
Russia (in Europe)	66,000	41
Union of S. Africa	62,000	13,000
Austria	59,000	-
Colombia	30,000	-
Indo-China	22,000	22,000
France	19,000	4,000
Belgium	12,000	-
Spain	10,000	2,000
Spitzbergen	9,000	-
Japan	9,000	-
Holland	6,000	-
Other Countries	24,000	3,000

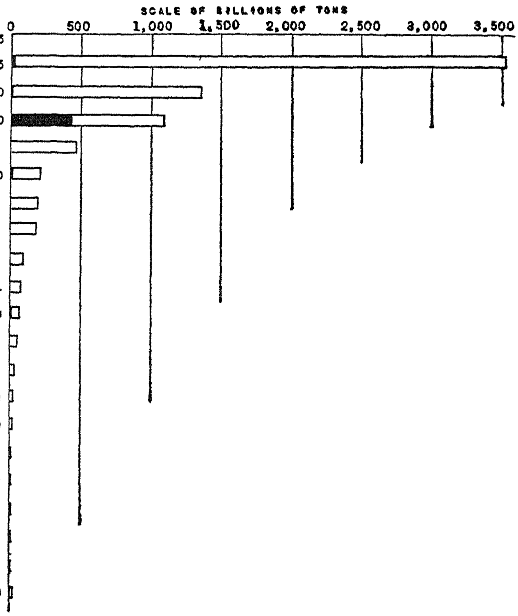


Fig. 97. Very Small Segments May Be Shown.

[1] *Cf.* Chapters L. and LI.

actual quantities, such as dollars, pounds, tons, or whatever the unit quantity be in which the data is expressed. It is an "absolute" chart, or chart of absolute values.

The other type of data is derived from the last, but is often more significant. In it the quantities have been turned into percentages, in each case of the totals for the line or stub. The totals therefore are in every case 100% and the bars are of uniform length—literally a series of 100% bars. The scale for such a chart is measured off in percentages instead of actual

PRE-WAR OCCUPATIONS
of the wage-earning population
IN SEVEN EUROPEAN NATIONS

	Agriculture	Commercial	Transportation	Mining, etc.	Manufacturing	Construction	Textiles, Dress	Miscellaneous
England	12.7	11.4	8.2	5.0	7.9	6.8	14.2	32.8
Belgium	21.9	11.8	2.0	6.6	6.0	7.3	14.7	29.8
Germany	35.1	6.3	2.9	3.3	7.0	7.0	9.1	29.3
France	41.6	6.5	2.9	1.6	4.4	4.2	12.6	26.4
Italy	59.1	3.4	3.1	0.9	2.1	5.0	11.4	14.9
Austria	60.8	3.3	1.7	1.6	2.8	3.0	7.2	19.6
Hungary	70.1	2.6	1.5	0.8	2.2	1.5	3.2	18.1

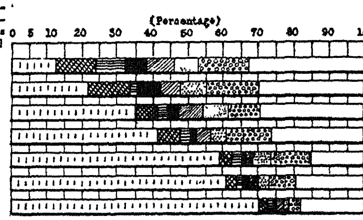
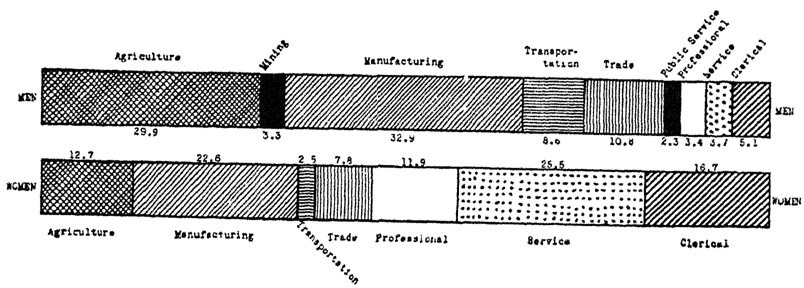


Fig. 98. The Relative (or Percentage) Bar-chart.

values, and the chart is a "relative" chart, or chart of relative values. Where both absolute and relative charts are being shown together, it is a good practice to use black for the absolute data and bars and red for the relative data and bars, a



OCCUPATIONS OF THE EMPLOYED
Percentages of Population 10 Years Old and Over Engaged in Gainful Occupations
United States
1920
(Source - United States Census)

Fig. 99. Any Pair of 100% Bars Really Form a Relative Bar-chart.

practice to which the ordinary two-color typewriter ribbon and carbon papers easily lend themselves.

A very different kind of chart is the one for which we suggest the name of multiple bar-chart. Here, two or more series of values, which are in themselves all totals, have been interlarded or dovetailed and fitted into each other and the bars

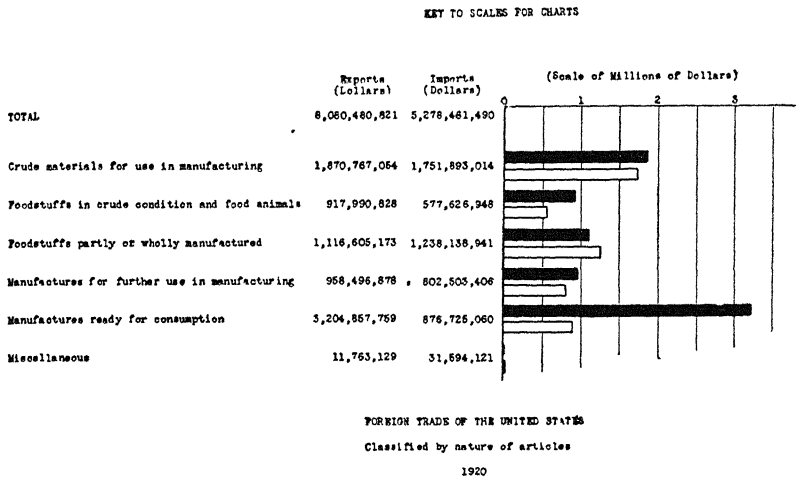


Fig. 100. The Multiple Bar-chart.

should be carefully and clearly distinguished in color and shading to show to which series they belong. The bars should generally be made narrower, to fit them closely together, into little groups, one group for each stub in the data. The data on the chart can still be prepared in distinct columns, although

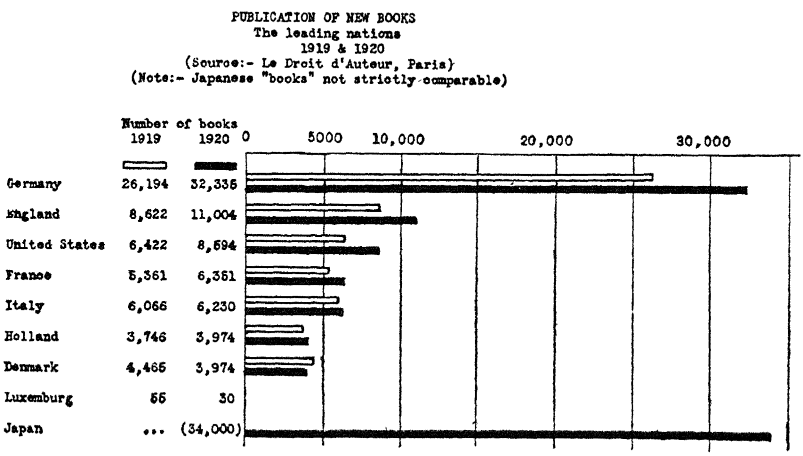


Fig. 101. A Good Comparison of Historical Data.

the bars have been interlarded, for the data is more easily consulted when not interlarded but kept in distinct columns. It is really a case of bringing two or more sets of bars together for cross-comparison. The result is rarely entirely satisfactory, and at best is confined to the combination of two sets, three or four sets being hard to make mutually distinct, and becoming confusing to the reader.

In fact, it can be laid down as a general rule that both the compound and the multiple bar-charts are too elaborate and complicated. A chart is always better the simpler it is, and we should make strong efforts to simplify these charts, and if possible reduce them to simple bar-charts. It usually pays well for sacrifices we make in this way, in legibility and interest to the reader, and after all, the chart of this type is generally directed at a reader, rather than at the maker. The only one

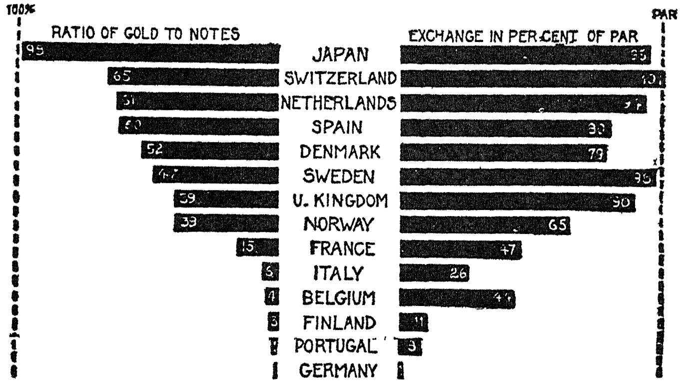


Fig. 102. Correlation is Indicated by Mirroring.

Ratio of gold reserves of Central Banks to paper currency in circulation compared with the relation of exchange rates to par value (March, 1922).—*Permission of Mr. Carl Snyder.*

of the three which stands out as absolutely simple and clear is the relative compound bar-chart, which consists of nothing more than a series of 100% bars.

The reason for the simplicity of the relative compound bar-chart is to be found in its uniform length of lines or bars. Only the segmentation of the bars changes in the chart, and the reader is not called upon to judge at the same time of various lengths and parts, but only of the various parts. In order to secure something of the same symmetry that marks the relative form, many chart-markers prepare the absolute form of

compound or segmented bar-chart in a pyramidal or bi-lateral form, the individual bars being aligned, not with even left hand ends, but with even centers, and each bar extending equally to right and left of the center line down the middle of

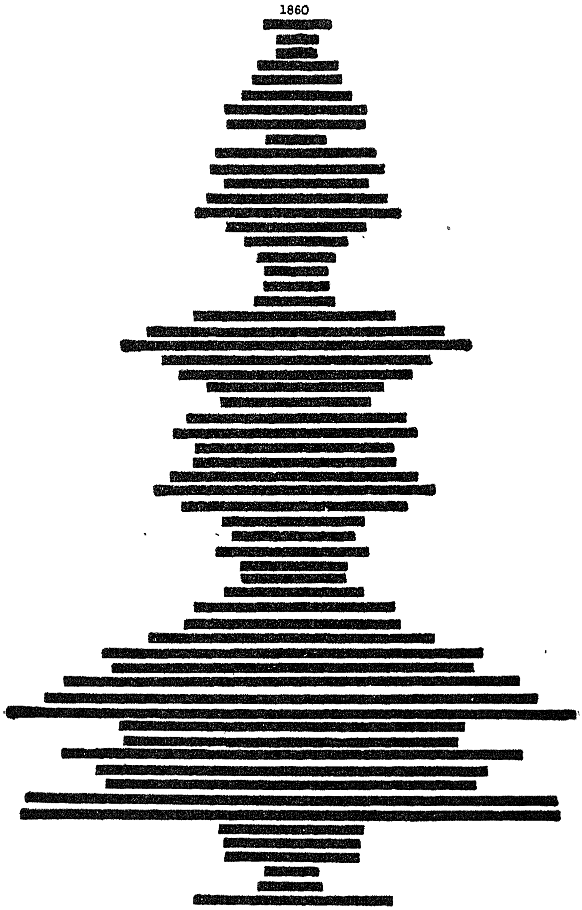
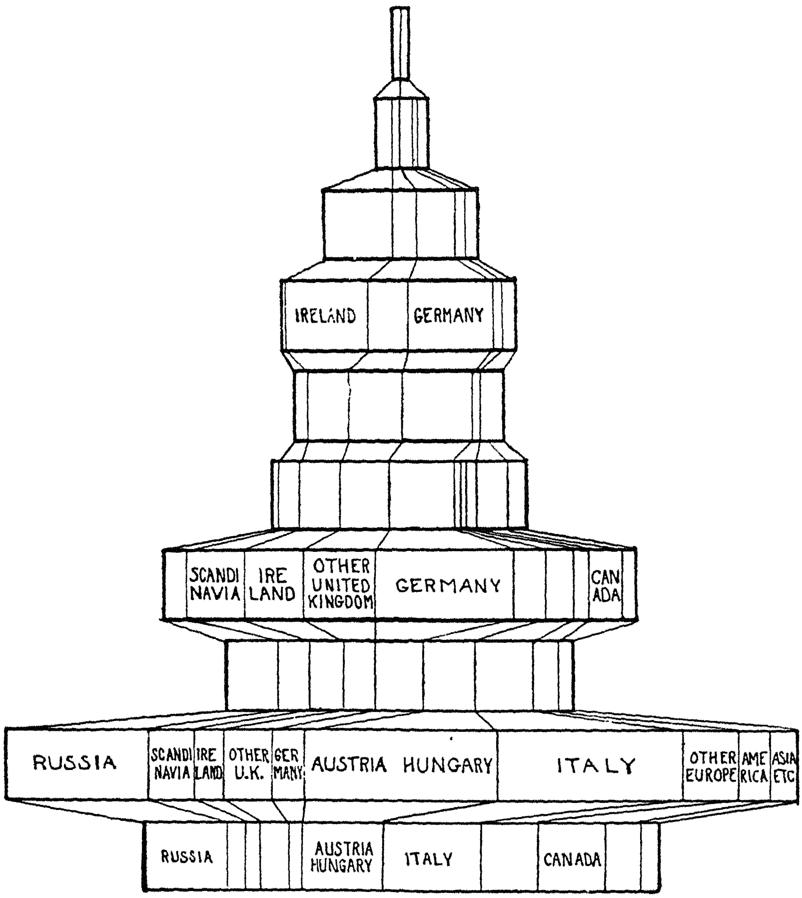


Fig. 103. Symmetry has Only a Popular Value.

the chart. Even simple or unsegmented bar-charts are sometimes arranged in this way. The form is certainly more pictorial and decorative than the forms which have been described. But data is, of course, not easily attached to this chart; in

fact it is ordinarily omitted. When plenty of space for the chart is available, and data is not going to be shown, but extremely pictorial or sensational effects are desired, this symmetrical form would appear to be entirely permissible. But it is certainly not to be recommended under any other conditions.

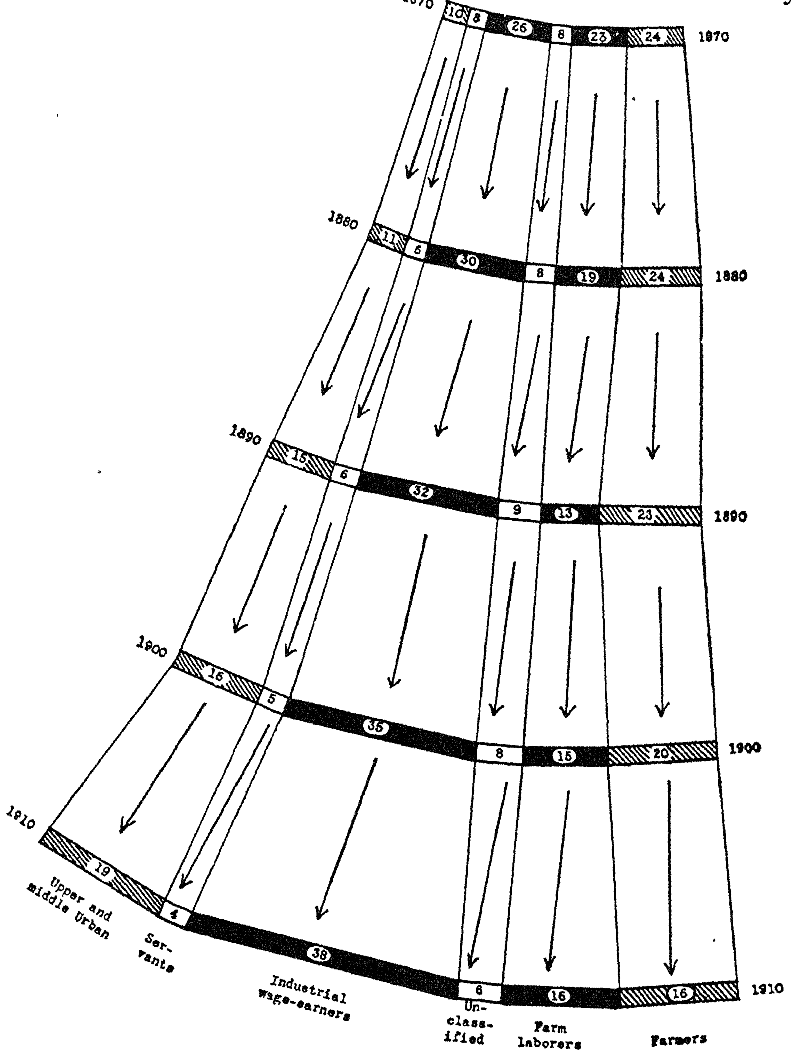


A CENTURY OF IMMIGRATION
Countries of last permanent residence of Immigrants Arrived
United States
1820-1920, by decades

Fig. 104. Connection Lines or Shadings to Distinguish Segments.

A common practice in the elaboration or decoration of bar-charts both simple and compound, particularly when they represent connected data, is to draw connecting lines across

intervening spaces between corresponding division points on adjacent bars. In later chapters this process is dealt with fully, and the omission of bars described, but we may here



CLASS ALIGNMENTS OF POPULATION
United States
1870-1910, by decades
(from data derived from Census by A. H. Hansen)
(Note-- Figures show percentages, bars show quantities)

Fig. 105. Warping the Chart to Show the Trend of Changes.

observe the effects where the bars are kept and the connecting lines drawn in. The reader is, of course, enabled to identify and compare corresponding segments of the bars more easily. In an extreme form of this chart, sometimes called the "stream

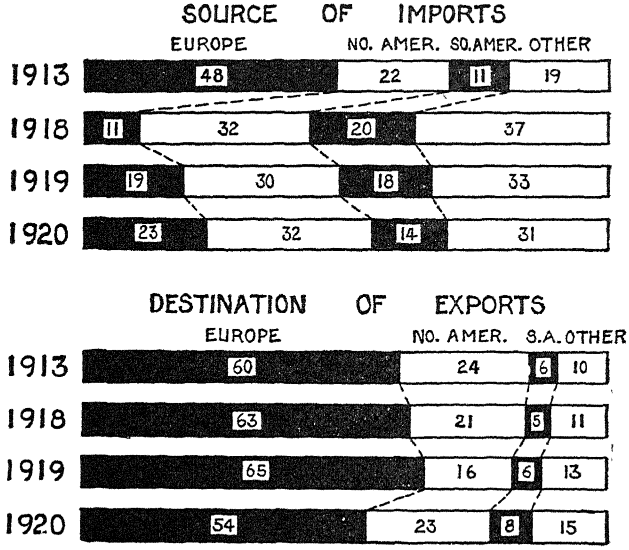


Fig. 106. Connection Lines. Note Inserted Data.

Percentage of total imports received from different continents, and percentage of total exports shipped to different continents.—*Permission of Mr. Carl Snyder.*

chart," the bars are broken or bent so that corresponding segments are kept as close as possible together. The chart is adapted only for data in which the changes of segments are fairly uniform and the result is entirely popular in its appeal, having no research value at all. What will be later described as a smoothing process, takes place in the segmenting division lines across the bars, and they, together with their connection lines, are made as nearly as possible straight rather than zigzag or rectilinear lines. There is little to be said for either the symmetrical bar-chart, the stream bar-chart, or the connection lines between the bars, but they are here described as examples of the modification and variation to which the bar-chart itself is susceptible.

It would, of course, be possible to go on and still further combine compound and multiple bars until we have a compound

BOOKS PUBLISHED IN THE UNITED STATES AND ENGLAND

New books, new editions and pamphlets published in 1920 compared as to subject, for the two countries.
(Source:- *The Publishers' Weekly, N. Y.*)

(American in plain type)
(British in italic type)

	New Books	New Editions	Pam-phlets	Total
TOTAL	5,101	1,086	2,235	8,422
	7,675	*2,266*	*1,063*	*11,004*
Philosophy	209	33	32	274
	233	*31*	*12*	*276*
Religion	467	37	161	665
	573	*77*	*29*	*679*
Sociology	353	43	363	759
	687	*61*	*222*	*870*
Law	70	39	57	166
	213	*76*	*74*	*363*
Education	101	10	123	234
	164	*17*	*72*	*253*
Philology	141	54	49	244
	176	*26*	*3*	*205*
Science	182	49	281	512
	411	*94*	*92*	*597*
Technology	259	93	183	535
	437	*155*	*128*	*720*
Medicine, Health	132	75	83	290
	269	*123*	*54*	*446*
Agriculture	49	18	223	290
	146	*33*	*39*	*218*
Domestic Economy	22	6	21	49
	57	*13*	*3*	*73*
Business	144	24	78	246
	103	*16*	*19*	*138*
Fine Arts	94	6	30	130
	158	*16*	*10*	*184*
Music (works about)	44	5	23	72
	59	*6*	-	*65*
Games and Sports	50	10	52	112
	120	*33*	*8*	*161*
Literature	248	53	50	351
	299	*52*	*15*	*366*
Poetry and Drama	409	44	105	558
	436	*80*	*47*	*563*
Fiction	778	345	31	1,154
	1,038	*1,051*	*15*	*2,104*
Juvenile	410	67	22	499
	613	*148*	*9*	*770*
History	503	36	172	711
	438	*44*	*43*	*525*
Geography, Travel	144	22	56	222
	424	*71*	*109*	*604*
Biography	271	14	29	314
	340	*33*	*1*	*374*
Miscellaneous	21	3	11	35
	181	-	-	*181*

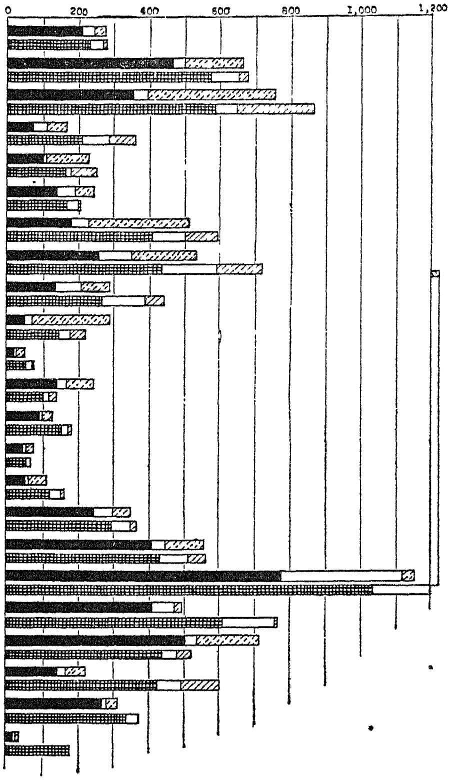


Fig. 107. A Compound Multiple (Absolute) Bar-chart.

multiple bar-chart. Rarely you may have use for such an animal, but it really lies out in that field of freak charts into which the enterprising chartist will inevitably wander by himself, and from which he will surely return if he keeps his senses. The field is wide open and there is unlimited opportunity for originality in the making and dressing up of bar-charts, but in the last analysis all that matters is to tell a story and tell it well. You will generally find that this object is best attained with the simpler, sounder methods which have been

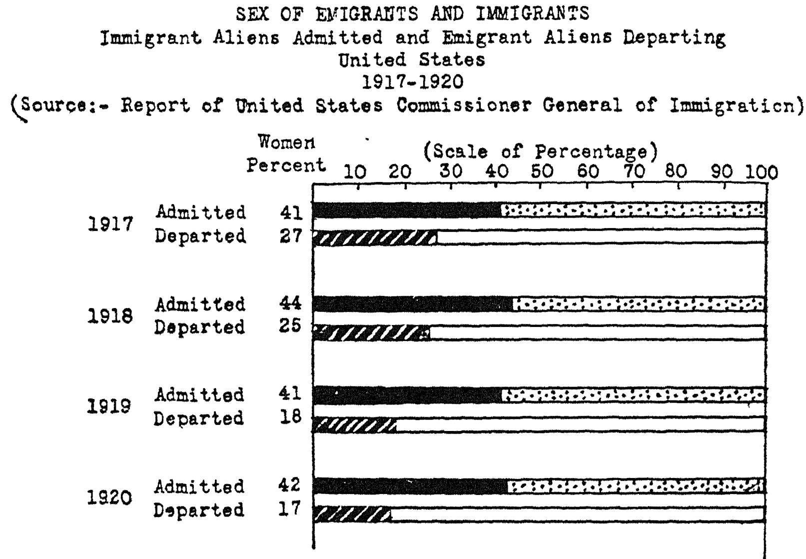


Fig. 108. A Compound Multiple (Relative) Bar-chart.

here discussed. In bar-charts, perhaps more than in any other form of chart-work, we must keep the purposes of simplicity and clearness always in mind, and avoid the more complex details which will suggest themselves, insidiously and attrac-

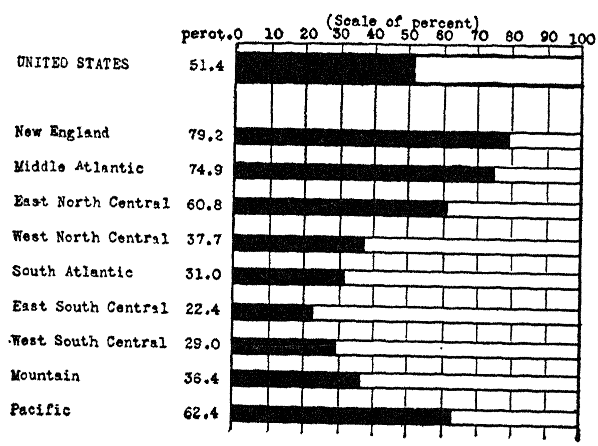


Fig. 109. The Compound Relative is the Best of the Composite Bar-charts.

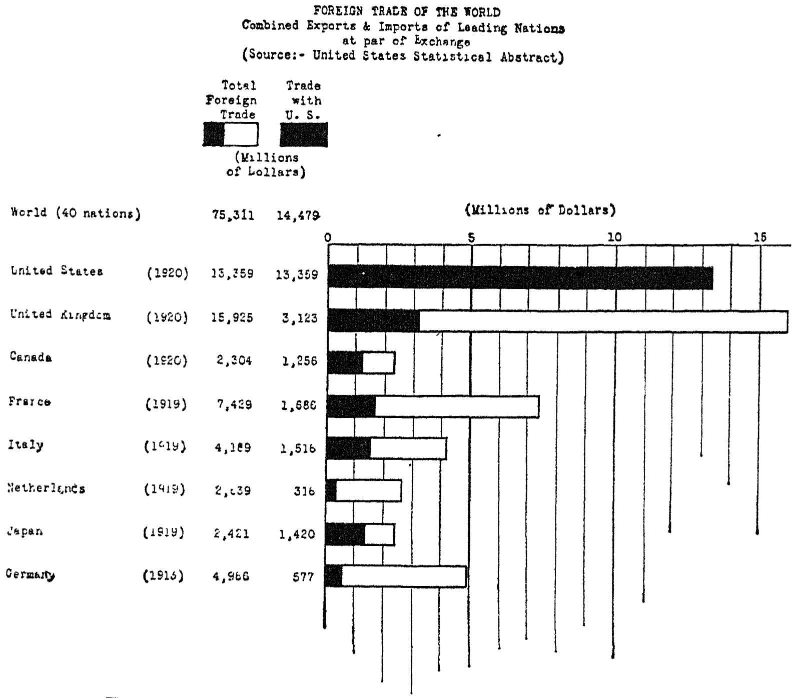


Fig. 110. The Simpler Forms Are More Effective.

tively to us as makers and designers. A simple chart which is read and understood is better than a complicated one which no one deciphers.

Chapter XIII

PICTORIAL BAR-CHARTS

For purposes of publicity, the circular form of chart has decided advantages over the rectilinear chart. This has been gone into in the chapter on pie-charts. The circle attracts the attention even of casual readers. And circular-shaped charts are therefore popular with all those propagandists who seek, by sugar-coating their information, to dispense it to an unwilling and indifferent public. The charts are useful for, and should be only designed for, advertising, and the popular presentation of educational matter. They are useless for research and study. These considerations have been discussed in the chapter on pie-charts, but arise again in connection with the possibilities of converting series of bars, that is, bar-charts, into series of circles.

Truly, when a series of circles are to be used in a chart, the chart-maker's road should be marked "Warning: Dangerous Curves Ahead." For the path he must pursue around the unshakable fact that a circular area has two dimensions, is at times devious and hard, in view of the rule against showing one-dimension data by two-dimension charts. The data which is shown upon bar-charts has but one dimension in the sense of the rule and the bar-chart itself has but one dimension. But when the bar-chart is converted into a series of circles, the result is extremely likely to have two varying dimensions and be as disastrous as the use of squares discussed in the chapter on dimensions.

We have seen that the substitution of a single circle for a single bar (or 100% bar) is harmless, for the reason that the area of the segments of the circle vary directly with the arcs and subtending angles. In short, in the pie-chart the areas of slices of the pie vary directly with the linear or one-dimension variations, and there is no conflict of measurements. The chart is like the 100% bar, for that too has an area which varies

directly with one of its linear measurements. In the bar the width is constant; in the circle the radius is constant.

But a series of bars of different lengths can not rightly be turned into a series of circles of different circumferences. In a series of bars the widths of the bars can be kept constant and hence their areas can be made to vary directly with their lengths. But in a series of circles of various circumferences the radii cannot be kept constant, but must vary with the circumferences, so that the areas of the circles will vary by the squares of the variations of the circumferences. In short, the moment you use circles of different sizes, the old conflict between area measurements and linear measurements creeps in, the most fundamental principle of charting is violated, and the chart becomes fallacious and deceptive.

There is but one type of bar-chart in which the bars can be turned into as many circles, and that is the relative compound bar-chart of the last chapter—a chart which is nothing more than a series of 100% bars. As neither the total length nor area of these bars varies, they can be safely turned into as many pie-charts or 100% circles of uniform circumference and area. The chart is one in which the segments alone are significant. It is true, as was said in the chapter on pie-charts, that the segments cannot be so well compared as in the relative compound bar-chart, and for this reason the chart is of less value for careful study, but the circular shapes have been secured and the chart has been perhaps made more attractive and popular.

For the simple and multiple bar-charts there is a dodge by which circles can be used, if you are intent on circles at any cost. The result is not appreciably less interesting and it has the advantage of being accurate. It consists in using whole circles and fractions or fragments of circles, all of uniform radii. Adopting one value in the series—perhaps the average—as 100%, you must turn all your data into percentages before preparing this chart and then plot as many circles and fractions of circles as the data calls for. This method of charting is sound because throughout the circles and fragments of circles, a uniformity of radii has been maintained, and the areas vary directly with the circumferences and arcs.

The drawing of these charts is comparatively easy, as all circles and parts of circles can be put in with a bow pen or compass without changing its setting. The work involved is

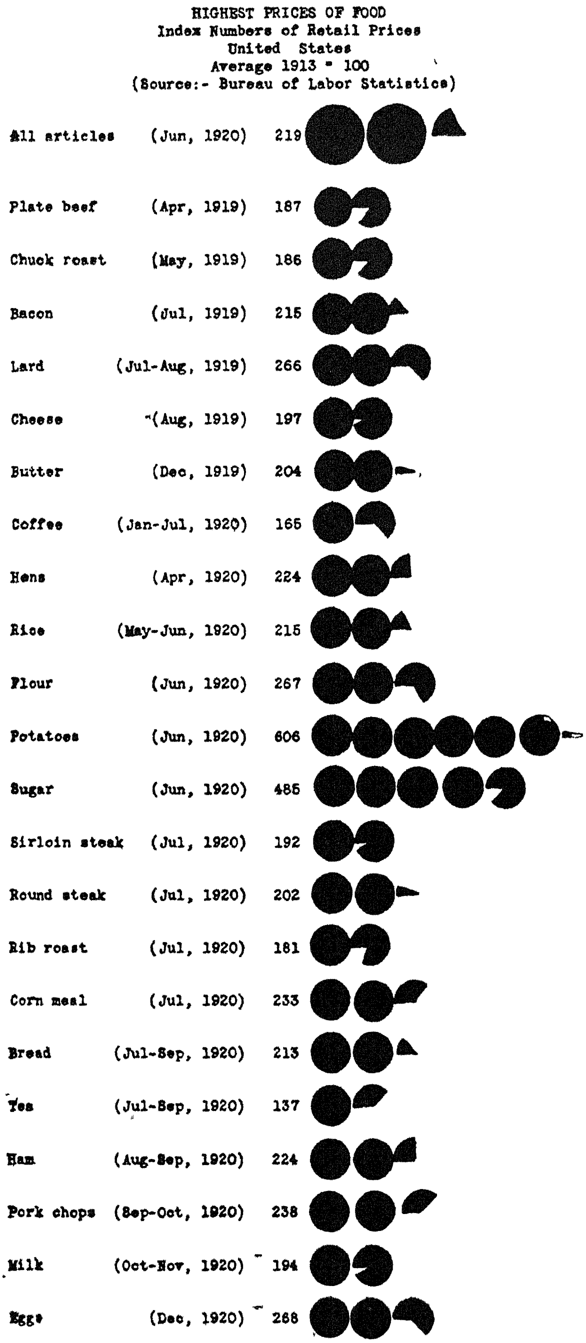


Fig. 111. The Circles Must Have Uniform Radii.

far less than it would be if circles of different radii had been used, requiring fresh setting of the pen for each circle,

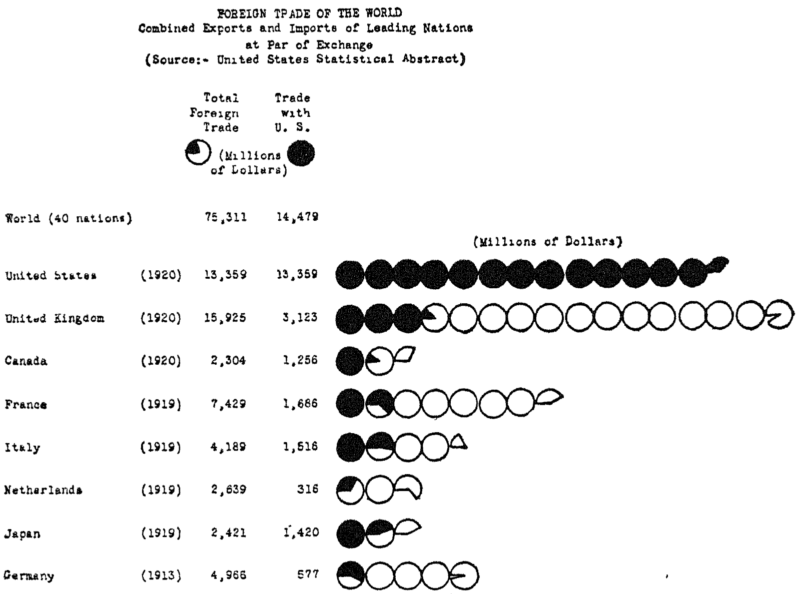


Fig. 112. Segmented, Like the Compound Bar-chart.

from square root calculations of the variations. The segments or fragments of circles can ordinarily be drawn in at

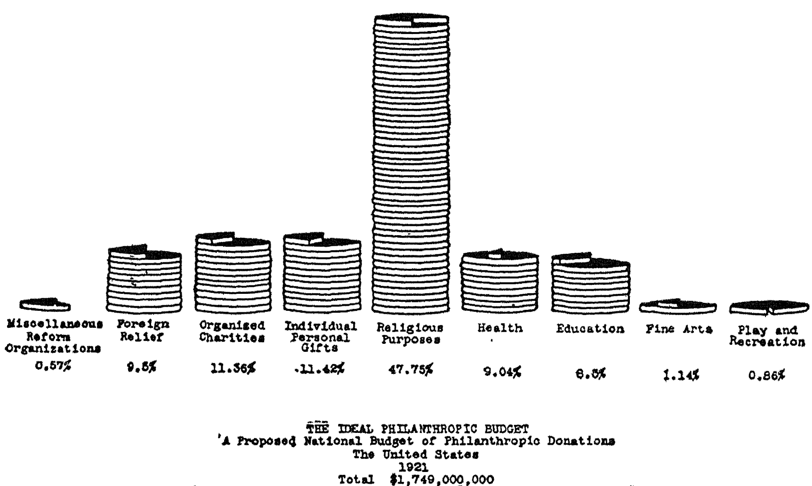


Fig. 113. Suggesting Metal Coins.

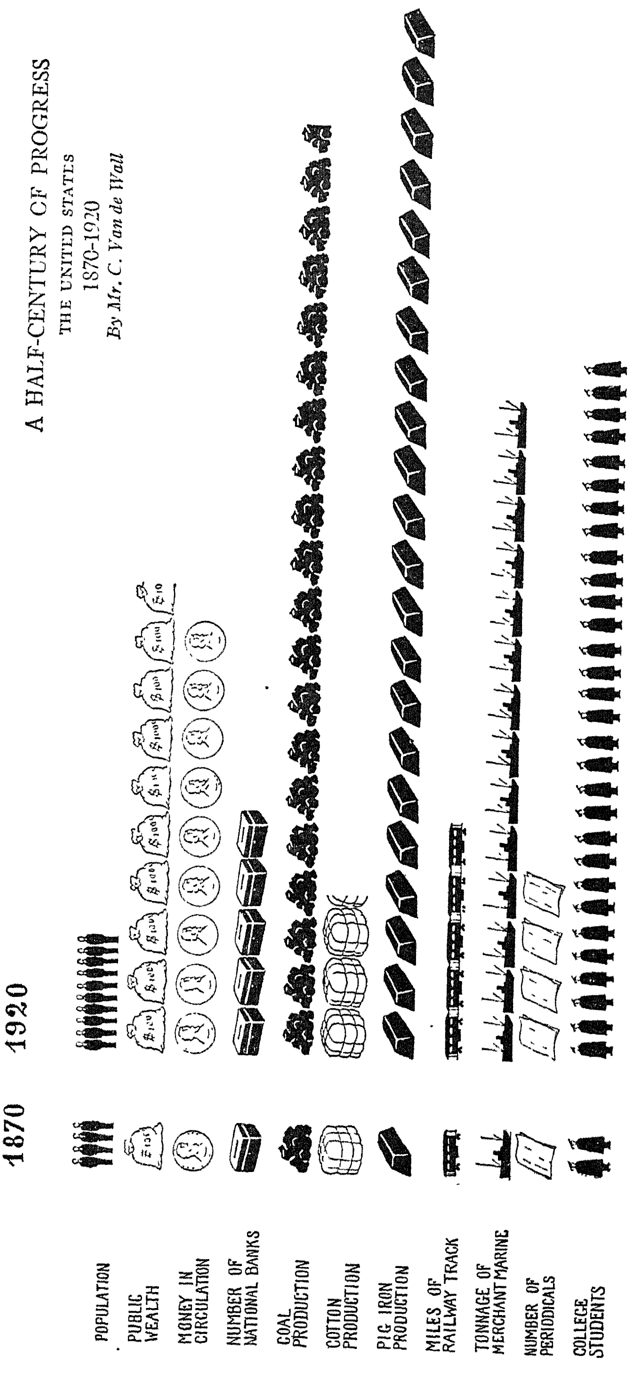
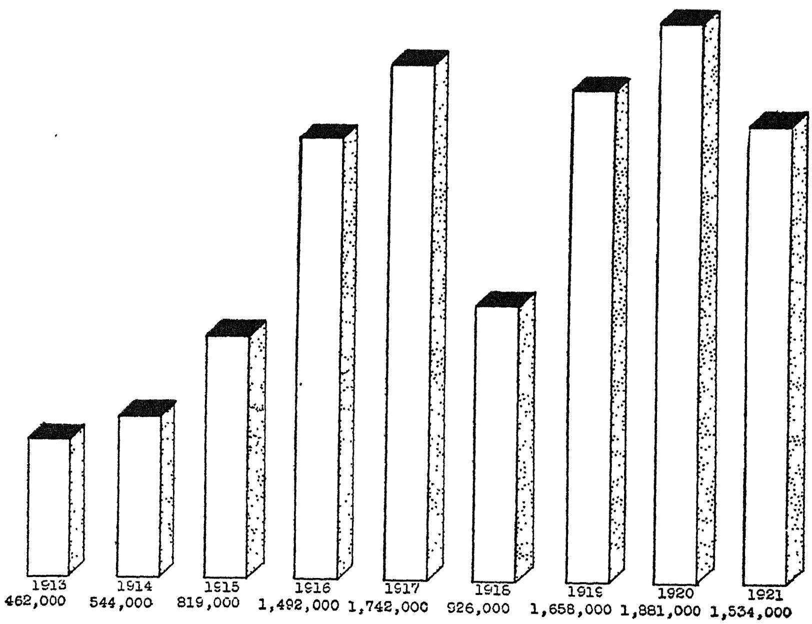


Fig. 114. Pictorial Figures May be Substituted for Bars.

sight, very accurate work for which protractors would be necessary not being of any value in this chart.

This form of chart is largely an attempt to present bar-chart information popularly. It is for that purpose particularly adapted to financial data, in which the circles can be taken to represent dollars and the fractions of circles parts of dollars. While in strict theory the circles should be at even distances from each other, yet where there are several for a single bar or figure, the conception of metal money is so vividly presented that the circles can be overlapped. The overlapping or circles saves much space without lessening greatly the impression on the reader's mind. It is as if, in the West where silver dollars are still used, you should lay out a row of these coins, each, except the first, tilted up and resting partly on the next one. But where space does not require this crowding up of circles, it is better as a general rule to place them at



AUTOMOBILE PRODUCTION
Number of Passenger Cars Produced
The United States
1913-1921
(Source:- National Automobile Chamber of Commerce)

Fig. 115. The Third Dimension Is Ornamental.

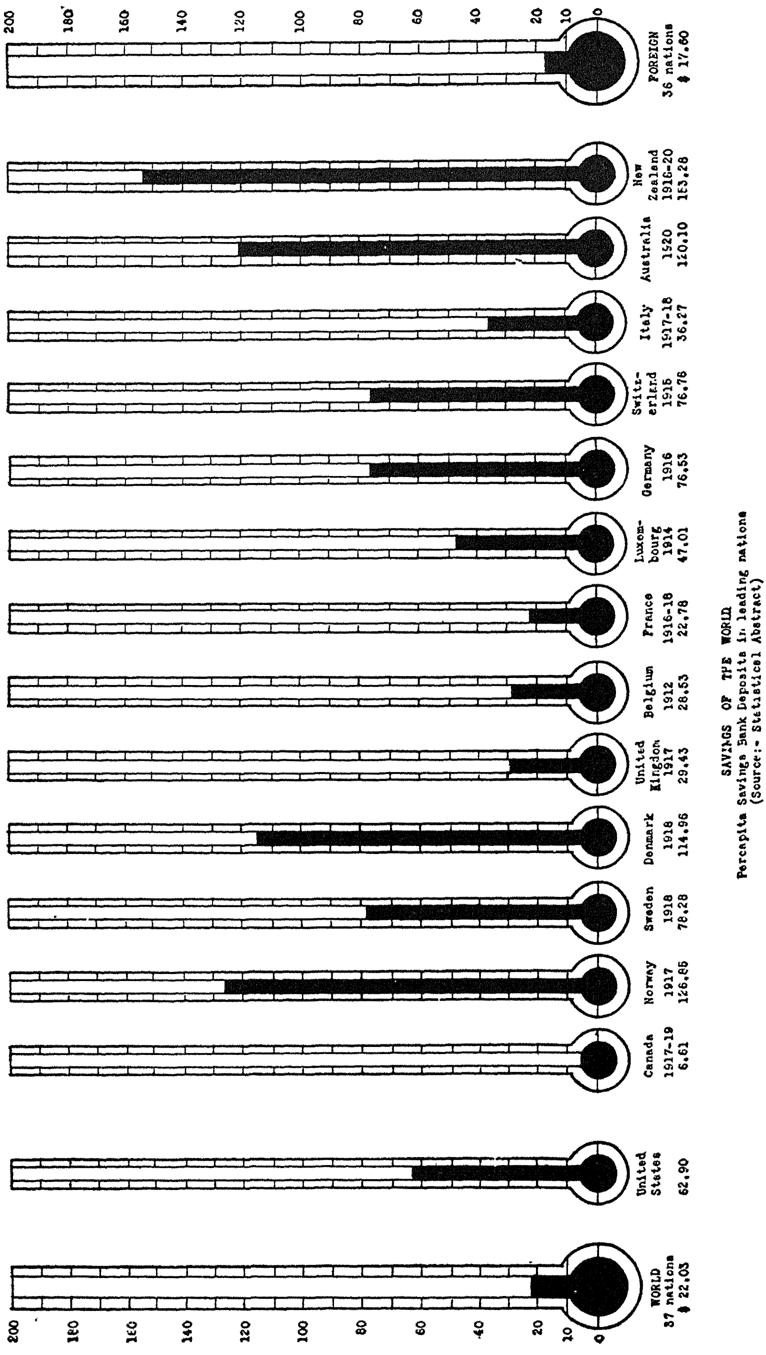
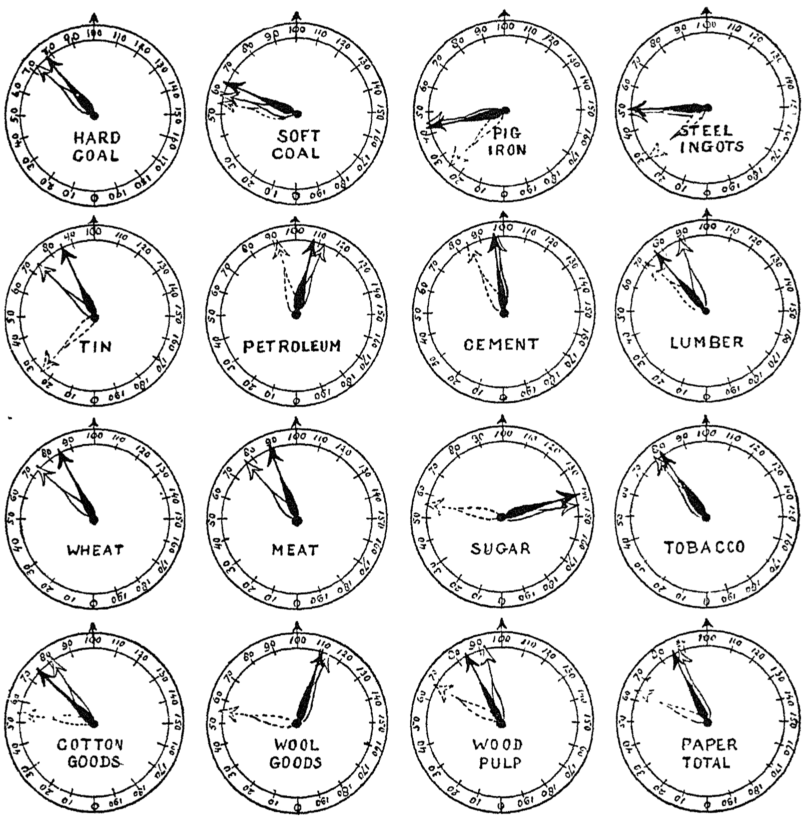


Fig. 116. One of Many Devices to Stimulate Interest.

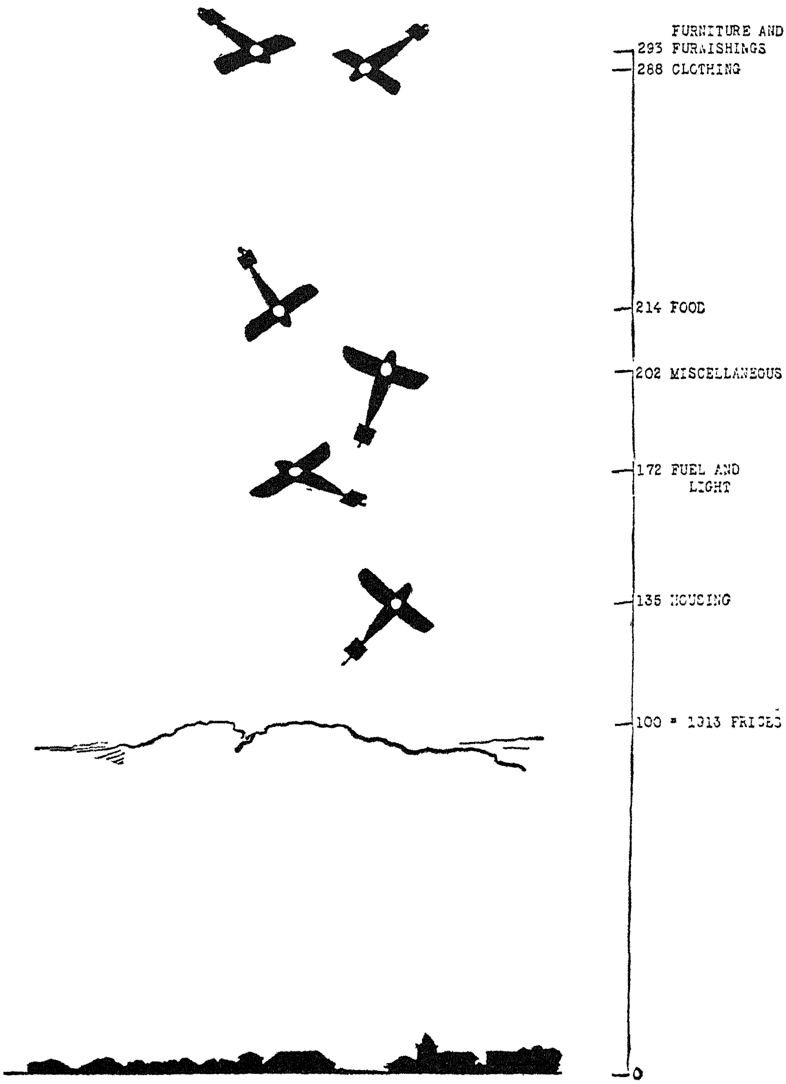
even distances. They will then roughly form bars of circles or coins.

Nor do you need to present a row of circles in the place of your bars. Rows of human figures, all drawn to the same scale, can be used in the place of bars. The length of the rows and the number of figures depicted in each, will show quite as well as plain bars would have shown, the various amounts represented. This is a method which those who wish to com-



PRODUCTION OF BASIC COMMODITIES
Index Figures of Monthly Production in Specified Industries
United States
Dec. 1921-Jan. 1922
(Source - Federal Reserve Bulletin)
(Normal = Trend after Allowing for Seasonal Variations and year-to-year growth = 100%)
(Black Arrow = Jan. 1922)
(White Arrow = Dec. 1921)
(Dotted Arrow = Low of 1921)

Fig. 117. Pointers, Instead of Segments, Suggest Pressure Gauge Dials.



THE HIGH COST OF LIVING
Index Figures of Retail Prices
United States
June, 1920
(Source:- Monthly Labor Review)
(1913 Average = 100)

Fig. 118. Aeroplanes, Horse-races, Boat-races, and the Like, Have a Certain Popular Value.

pare pictorially two or more populations, can safely employ. Instead of showing the Japanese army with a single small soldier and the American army with a single large one, in which case you confront the reader with three conflicting measurements—height, surface area, and cubic volume or weight—you need merely show one Japanese soldier and several American ones, all of the same size, and their number will give an accurate conception of the relative sizes of the two armies.

Here, then, is the answer to the problems raised in the chapter on charting principles. Here is the proper way to show pictorially the comparison between two or more items. Do not draw one loaf of bread and an enlarged replica of it beside it, to show how much the food-bill of the nation has changed, but draw one loaf of bread and label it with the earlier year and draw several loaves of the same size and label them for today. Do not bring together a large and a small house nor a large and a small nugget of gold, nor a large and a small railroad car, but place together a single one of each and a group. The number of times this simple rule is violated, with results which vary between gross understatement of a perfectly good case and gross deception about a poor one, will amaze you when you begin to watch for it. And the amount of money spent sometimes in publishing them, futile or false as they are, will also amaze you. It is one of the most frequent of all errors in charting.

In addition to the geometric pattern of the rectangle and circle, there are countless pictorial devices, the simplest of which is to indicate a third dimension to the bars, setting them up on end for this purpose. The ingenuity of advertising artists has hardly been tapped as yet, and thermometers, barometers, or pressure-gauge dials are but the beginning of the avalanche. The pictures of motor, horse, or boat-races, or altitude flights of aeroplanes have already been found useful, and it is probable that all popular contests can be made to yield attractive pictorial substitutes for the prosaic bar-chart.

Chapter XIV

VERTICAL-BAR CHARTS

Between the sensational picture-bar which we have just considered, and the plain bar itself, there is a type of bar which is both accurate and popular. Of less value in the statistical laboratory, it nevertheless deserves a passing glance even from the most academic investigator, for it forms an interesting link between bar-charts and higher things.

To make a bar-chart popular, knock it over flat on its side, so that the bars stand up on end. Simple, isn't it? But that's

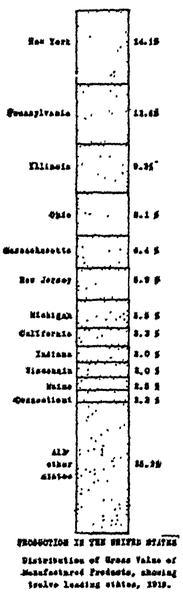


Fig. 119.

the rule. There being nothing more to discuss in the matter of making popular bar-charts, we are tempted to close the discussion at this point and produce a pleasant surprise to all.

But the vertical bar-chart is rich in suggestions for the higher forms of charts which we are approaching, and it deserves a close study.

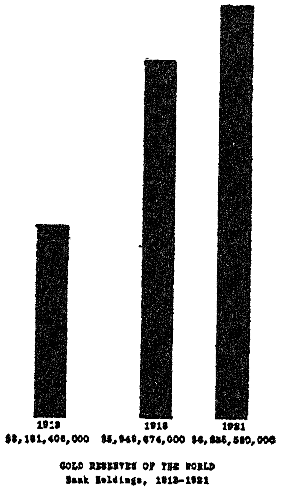
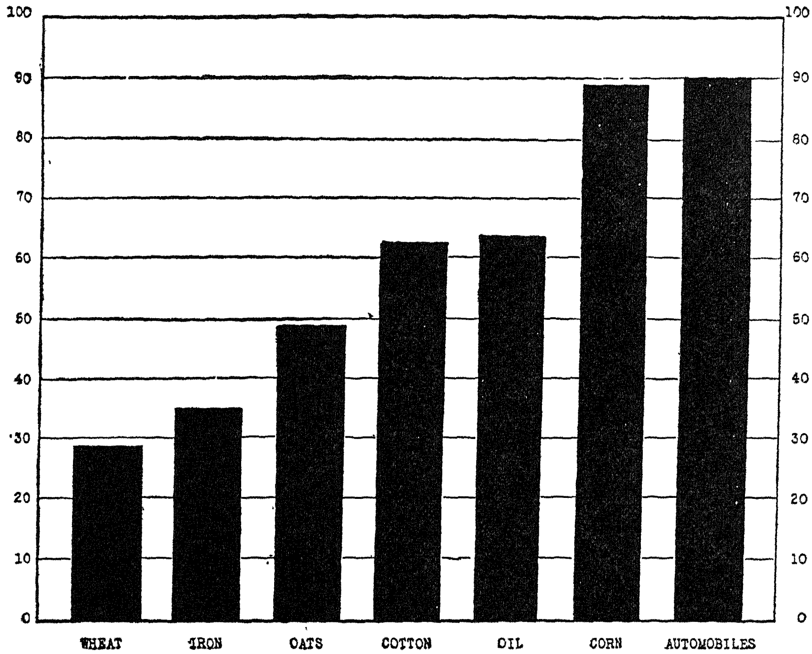


Fig. 120.

The chief value of the "pipe-organ chart" as it is sometimes called, lies in the realistic picture it gives of quantities. From a base line these quantities are seen to rise the full length of the bars, as so much substantial material stacked neatly in piles where we can compare them. We view them from the level or floor on which they are piled. We do not have to climb up and get a bird's-eye view of them as in the ordinary bar-chart, where we seem to be looking down upon rows and rows of goods, but we see them from a natural view-point. Nor do we rely upon an arbitrary arrangement by which their left ends have been brought together as in the bar-chart, but we know instantly that if they are piled up, it is their tops which we must watch. The pipe-organ chart finds instant response in our minds, and appeals to us as both logical and natural. A child can comprehend it.

If you call this base-line the x-axis of your paper and give the upright bars values in the y-axis, you will be reminded of co-ordinates and maps. But it is not necessary to go so far. Merely think of your own back yard, and the nice high fence about it which you have just white-washed. Assume that through some weird freak of carpentry you built it with



UNITED STATES PERCENTAGE OF THE WORLD'S PRODUCTION
of specified commodities

Fig. 121.

boards which run horizontally. Or turn and look at the wall of your house, with the weather-boarding running horizontally about it. Against such a wall let us pile your quantities in neat columns or let us stand up some dark boards of the right height against it. You are then ready to take a photograph which will be a good pipe-organ chart. The lines of the weather-boarding on the house will make the field of the chart, and the upright dark boards will be the bars.

Note also, and this is important, that if through standing too close you should take a picture showing only the upper ends of the upright boards, but not their full lengths, you would consider the resulting picture not only a failure but actually deceptive. In other words, you must not omit the zero-line or base-line. While you would succeed in showing the variation of the top ends more clearly you would no longer have comparable lengths. One board might be but a tenth longer than the other, but by cutting the lower eight-tenths

out of your picture, it would appear to be twice as long. The thing simply could not be done, unless you wilfully undertook to deceive yourself or someone else. The conception of the pipe-organ chart is sound and fundamental. It is perhaps the most direct charting method we have. It is almost fool-proof, which is more than can be said of most charts. And it establishes clearly the vital principle not to omit a zero-line.

Moreover in the pipe-organ or vertical-bar chart, we first encounter labelling or data difficulties. And if there is one motto which we should like to print at the bottom of every page in bold-face type, as do the publishers of other valuable reference-books, it is this: "Never separate your chart from its data." On the contrary, incorporate the data in the chart. For a chart without its data is a poor lost thing indeed. And the unhappy reader wishing to know what it means must hunt

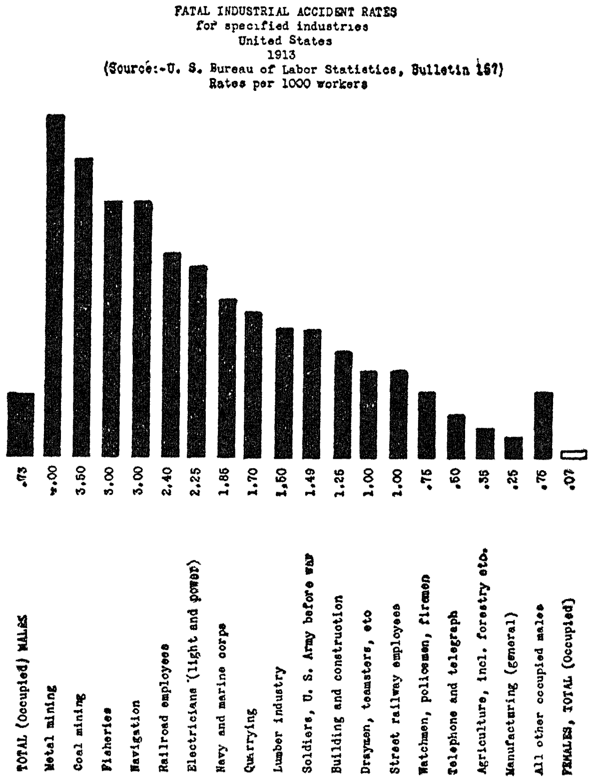
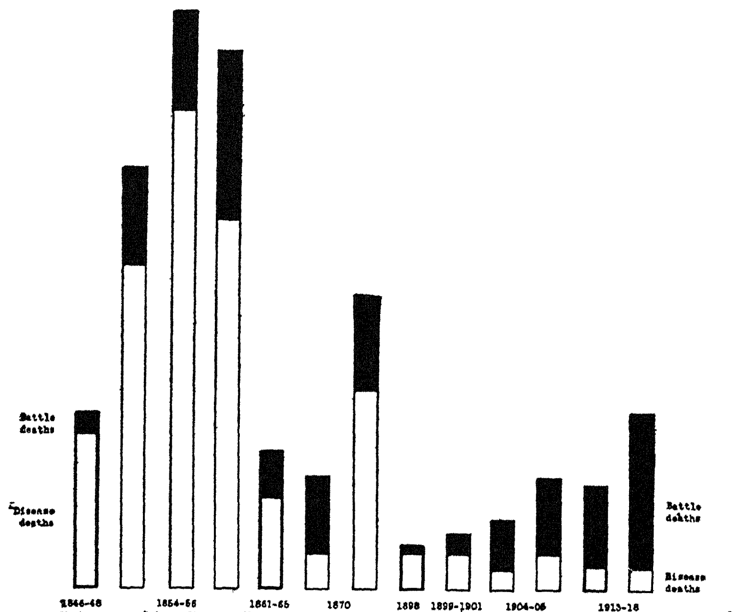


Fig. 122.

and hunt and hunt till he locates the particular information in some distant table. As a matter of fact, he won't do it, for before he has found his data he has lost his interest in the matter, and then what good is your chart?

In the pipe-organ chart, however, it becomes difficult to append data directly to the bars. Following your rule of tipping the horizontal bar-chart on end, you would naturally have the data down below the bars, reading upward laterally. This is at once a logical and a sound place, for the bars should be in line with their own data. But because the vertical bar-chart is for popular consumption, and because the average man does not care to crane his neck to one side and read on edge, objection is often raised to this method of disposing of the data.



Date War	1846-48 Mexican War	1854-56 Crimean War			1861-65 Civil War	1870 Franco-Prussian		1898 Spanish-American	1899-1901 Anglo-Boer	1904-05 Russo-Japanese		1913-18 World War		Date War
Troops	United States	British	French	Russian	Northern U. S.	German	French	United States	British	Russian	Japanese	U.S.A. A.E.F.	British E.F.	Troops
Total Death Rate	125	299	411	383	98	80	209	31	40	49	79	74	126	Total Death Rate
Disease Death Rate	110	230	341	263	65	26	141	26	26	14	25	17	16	Disease Death Rate
Battle Death Rate	15	69	70	120	33	55	68	5	14	35	54	57	110	Battle Death Rate

DEATH RATES IN WARFARE
Battle and Disease Death Rates per 1000 Soldiers per Year
In Specified Wars
1846-1918
(Source.- The Official United States Bulletin)

Fig. 123.

Nevertheless the method remains proper, and one is almost tempted to say, let the average man learn to crane his neck if he wants to check up on our plotting. As a matter of fact, the average reader is generally satisfied to know that the data is there where he can get it if he wants it, and so does not bother to look at it anyway. An occasional figure in which he

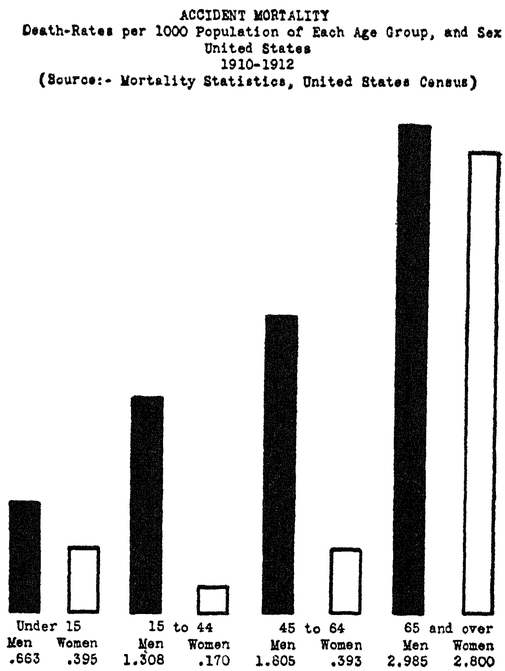
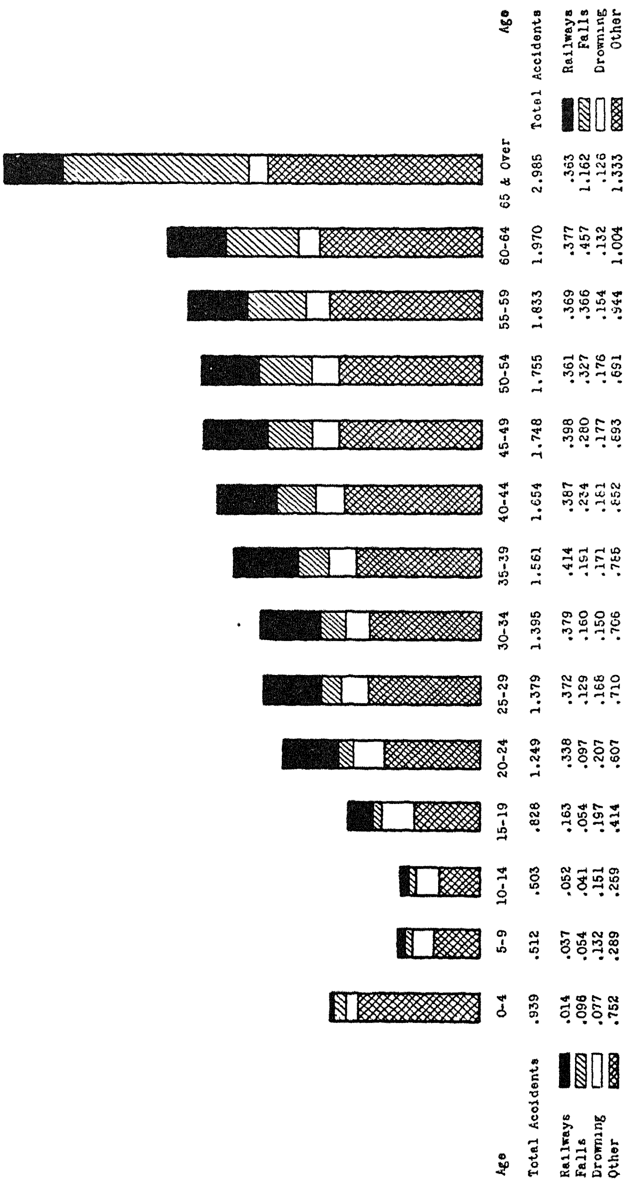


Fig. 124.

is really interested will be read carefully by him in spite of its reading upward, never fear. And throughout the whole field of charts it is of such great value to be able to place one's data or figures out along projected lines from plotted bars, or points, that we must adopt the upward reading data in spite of its temporary strangeness. It is to be accepted and adopted as a proper feature of charting.

Where the bars are very wide, or the spaces between them wide, there may, it is true, be room in which to write the data horizontally, in little boxes below the charts. This method is wasteful of space and compresses words and figures confusingly,



MALE ACCIDENT MORTALITY RATES
Death-Rates per 1000 of Population of Males of Each Age
For Specified Accidents
United States
1910-1912
(Source:- Mortality Statistics, United States Census)

Fig. 125.

but it is a "very-simplest" method which you will sometimes want to use in presenting large and simple diagrams to school-children. If you have the space to give to it, it is perhaps the better method for extremely popular work, but it is not to be

generally used as it is not in the long run satisfactory even for the average popular chart.

Two principles can therefore be garnered from the pipe-organ chart, first that the base or zero-line should never be omitted, second that data should be kept with the chart re-

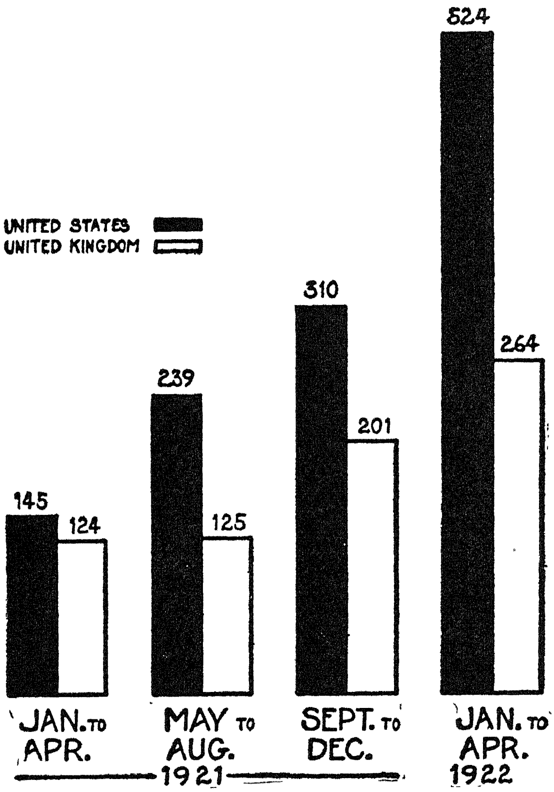


Fig. 126. An Absolute Multiple Bar-chart.
Volume of Foreign Financing in the United States and in the United Kingdom, in millions of dollars (pounds converted at current rates of exchange).—*Permission of Mr. Carl Snyder.*

gardless of the direction in which that data must be written. And as you progress further into charts, not only will it help you to retain these principles. but it will also help you immensely to visualize to yourself again and again this chart composed of vertical bars, a chart from which most of the higher forms have been evolved.

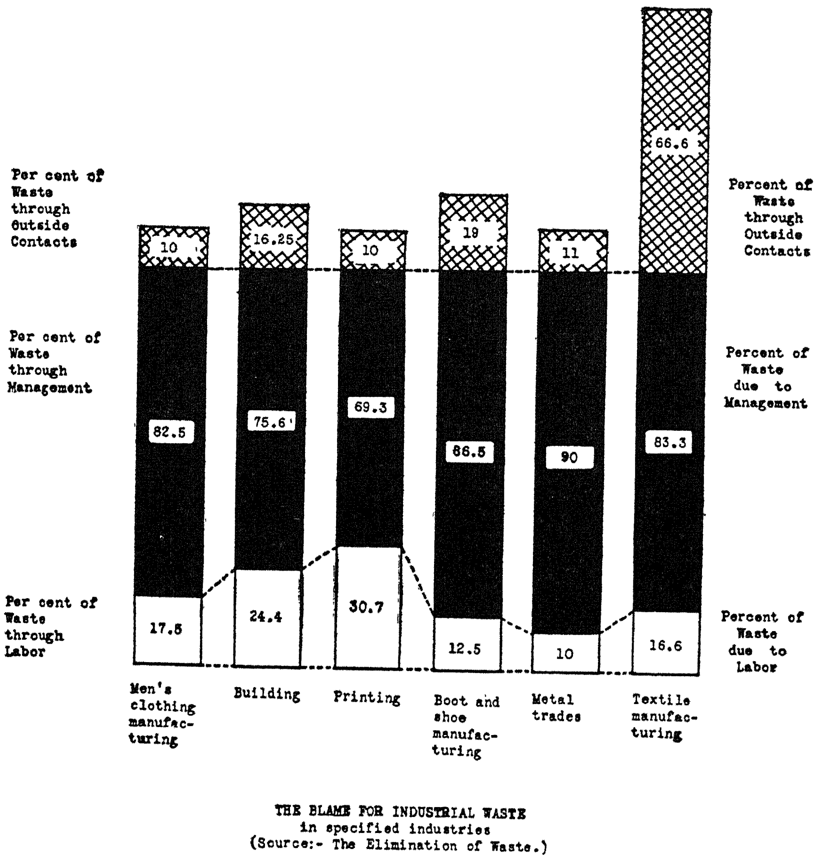
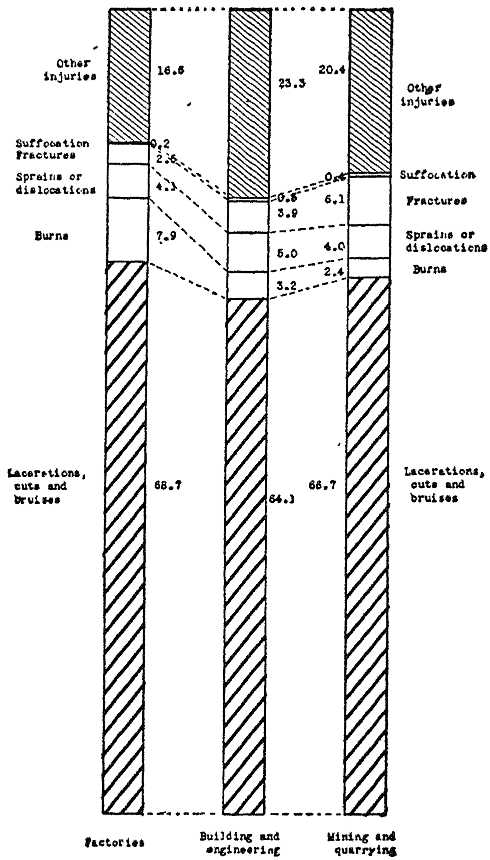


Fig. 127. Wide Bars with Data Inserted.



THE NATURE OF INDUSTRIAL ACCIDENTS
New York State
1911 - 1913

(Source:- N. Y. State Dept. of Labor)

Fig. 128. Connecting Lines Are Often Useful.

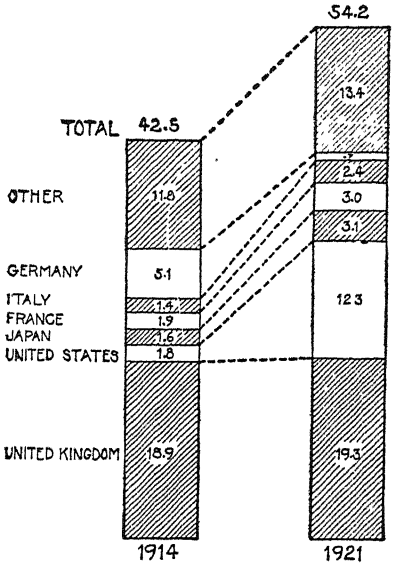


Fig. 129.

Gross tonnage of world seagoing iron and steel ships in 1914 and 1921 (in millions of tons).—*Permission of Mr. Carl Snyder.*

CHAPTER XV

CURVES

It may not have been a very clever fellow who invented curves, but he was assuredly lazy. For he balked at the task of drawing vertical bars in the pipe-organ style and he said, "Since I am only interested in the ends of the bars, I will place a dot where each bar ends, and let it go at that." And later when he wished to find the dots quickly, he drew connecting lines between them and, behold, he had a "curve." A curve can, therefore, be defined as a line passing through the upper ends of the bars in a vertical-bar chart.

1790	3,930,000
1800	5,310,000
1810	7,240,000
1820	9,640,000
1830	12,870,000
1840	17,070,000
1850	23,200,000
1860	31,400,000
1870	38,600,000
1880	50,160,000
1890	62,900,000
1900	76,000,000
1910	92,000,000
1920	105,700,000

THE POPULATION OF
THE UNITED STATES
1790-1920

Fig. 130. Here is the Data—Historical.

Let us step out into your back yard again and take another look at the upright boards which in the last chapter were left

standing against the wall of the house. Will it not be an excellent plan—if the house is not yours—to drive a nail into the wall above each board to mark its height? Then we can throw the boards away or let the children play with them. And if

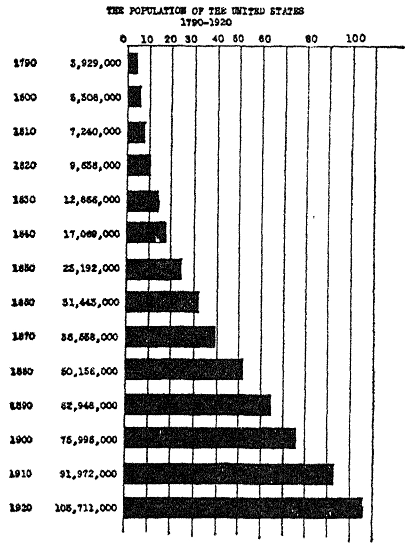


Fig. 131. The Ordinary Bar-chart.

we run a piece of dark string along from nail to nail, we will not have any difficulty in following the changes in their positions. Here we have a home-made curve. A photograph of this piece of string (as long as we also show the ground in the

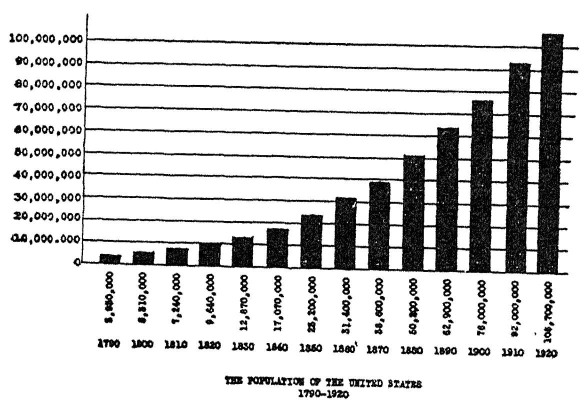


Fig. 132. Vertical Bars for Popularity.

picture) will do quite as well as a photograph of the original boards, for we can always imagine the boards running from ground to string. The picture will be complete if we run laths of uniform length up where the boards have been, that the exact position of the nails along the wall may be clear when we come to make the next curve on the same wall. In a chart these up-and-down laths which serve merely to mark the horizontal position of the now invisible bars, are called "ordinates" and their distances along the ground from the first lath are called "abscissae"—words never to be forgotten.

You will already have observed the wonderful thing about

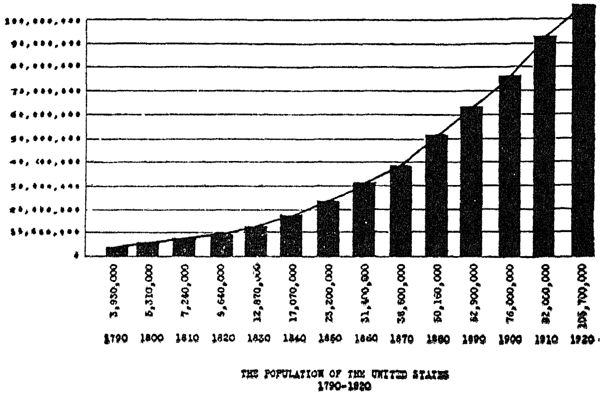


Fig. 133. A Curve Through the Bars.

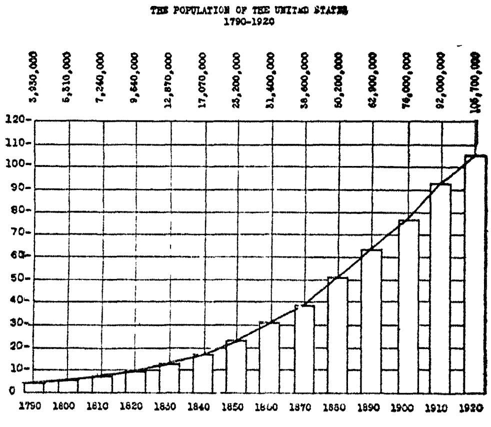


Fig. 134. The Bars Disappearing; the "Field" Appearing.

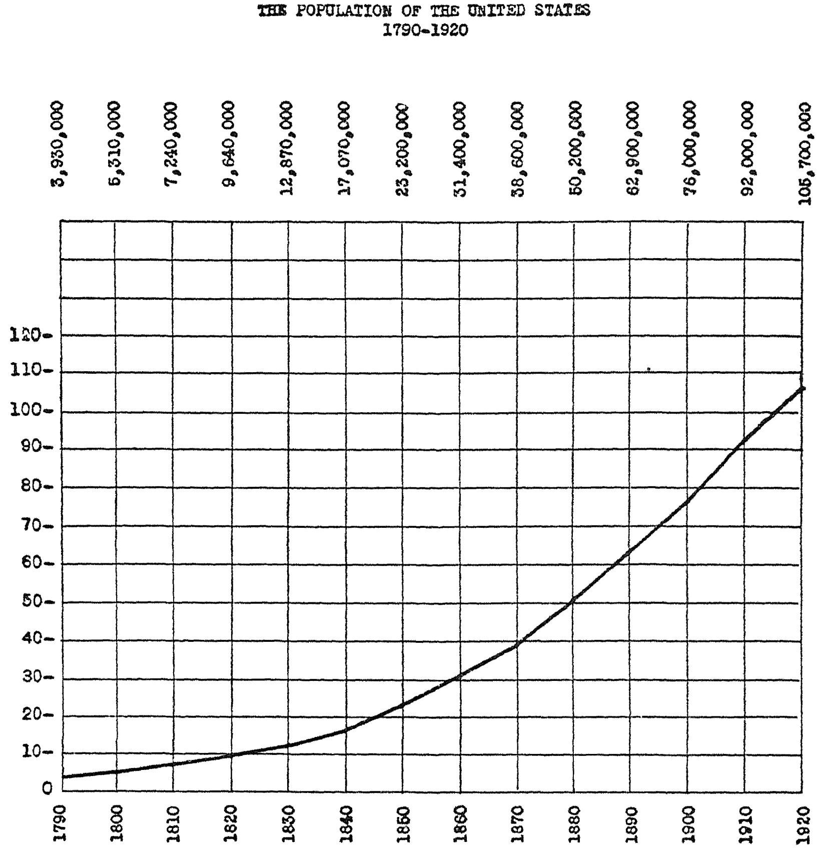


Fig. 135. The Evolution of the Curve is Complete.

a curve, namely that it is easily combined with several of its kind upon a single chart. Multiple curves are far better than multiple bar charts. A number of curves wiggling across the page at the tops of invisible bars are eminently more satisfactory than actual bars interlarded. In the first place, comparison of several series of data is greatly facilitated in curves because each set has been condensed and simplified into a single line. There is no difficulty in comparing values of each series with each other. In the second place, such a comparison is more accurate in curves because all similar points on various sets or series have been brought together upon a single vertical line. Had we placed the bars in this way on top of each other, the longest one would have wiped out or hidden all the shorter ones and in the multiple bar-chart, therefore, each set of bars

has to be shifted slightly out of position to avoid the next set. But when we plot only the end points of these bars, the short and the long ones show up equally clearly and can be brought together upon their true ordinate lines. In the third place a curve is much more easily drawn than a bar-chart. This is a most important reason, between ourselves. And, fourthly, to the reader who understands it (and it is a fact that schoolchildren understand it, however little the present generation of adults may) the curve is less confusing and more easily read for its salient points. You will find still other reasons why curves are advantageous as you go on.

A curve cannot, however, always be used in the place of a bar-chart, for the line which connects the various points implies that the data itself can be considered connected. Much data can not be so considered. A careful inspection of the data will soon show whether it is connected or not, for the stubs of connected data always form a variable. In the chapter on dimensions and variables, the test for variable nature in stubs was given somewhat as follows: "can the stubs or items be shifted up or down in their arrangement freely or is their order naturally fixed by their nature?" Variability is shown by the rigidity of order.

This limitation of the curve method can be made clear by two or three illustrations. We have before us the statistics of the population of the United States for each ten-year period during the last century. The stubs for each figure of population in this case are: 1790, 1800, 1810, 1820, 1830, 1840, and so on down to 1920. Now no sane person would think of arranging these normally in any order except from the earliest to the latest, or from the latest to the earliest—it would be ridiculous to adopt an arrangement such as the following: 1910, 1810, 1800, 1860, 1920 and so on. Clearly, this is a case in which the order of the items or stubs is naturally determined by the data itself and these various years can be considered as the various values of one variable, namely "time." The data can be charted on a curve. Consider another example. In taking a census of the buildings in a certain well-known village, the investigators returned reports of the number of one-story houses, the number of two-story houses, the number of three-story houses, and so on up to 55-story buildings. Here again the order of the items or stubs, that is the number of stories, is definitely fixed by the nature of these items and

the number of stories can be considered a variable in the same way as before and the data can be shown by a curve. Take

IMPORTS INTO RUSSIA
1921
(Source:- Russian Information and Review, London)

	($)
Total	124,281,000
Foodstuffs	16,061,000
Animal products	39,605,000
Timber and seed	504,000
Earthenware	227,000
Fuel, pitch, etc.	2,786,000
Chemicals	2,032,000
Metals, ores, machinery, tools	29,184,000
Paper and paper goods	3,977,000
Textiles	15,206,000
Wearing apparel, stationery, etc.	13,132,000
Miscellaneous (including 48,000 tons of "famine aid")	1,567,000

Fig. 136. Here is Data not in Series.

another example. The United States exported to England in the year 1920 a large amount of copper, wheat, rubber, automobile supplies, machinery and paper. If we were charting these exports, it makes no difference whether we show the cotton exports before the paper exports or vice versa. The order of these items is not fixed by their nature and can be arranged in any way we desire. Here, then, is a case in which we cannot use a curve but must fall back upon the bar-chart. In short, while the bar-chart can be used for all data, the curve-chart can only be used for data of which the stubs form values or readings of a mathematician's variable.[1]

[1] A curve or connected line in the place of vertical bars, for abstract or geographical data (that is, data of which the stubs are not an ordered numerical series) is a graphic monstrosity, fortunately not often seen.

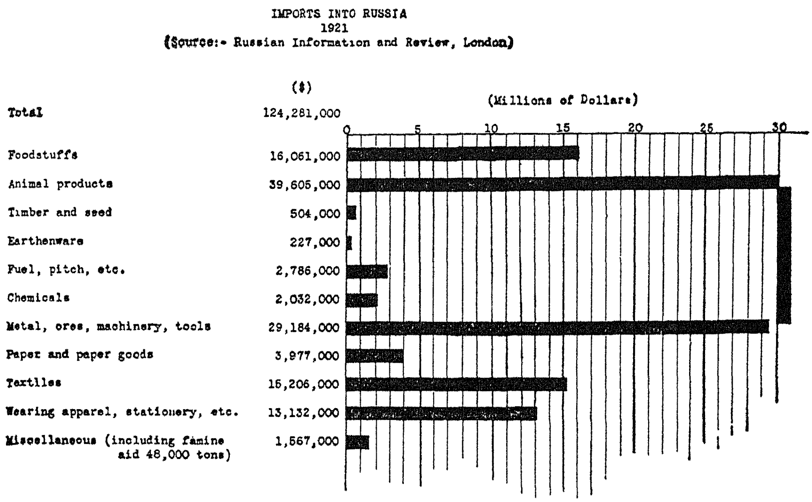


Fig. 137. No Curve Should be Made with This.

If we examine the field or background of a curve, we will find that it is drawn up according to the principles of Cartesian co-ordinates which we have already observed. The reader who has forgotten or omitted that weird chapter had best turn back to it and read it carefully. In order to understand co-ordinates for curve-chart work, you must know that the x-axis of your chart is that straight horizontal line along the bottom of the chart which we sometimes call the base line of the chart, or zero line. The y-axis is the vertical line whose value at all points is zero on the x-axis. In the ordinary chart, x and y axes are along the edges of the chart, the x-axis at the bottom, and the y-axis at the left hand side. In this case there is no room on the chart for negative values. It is not at all uncommon, however, to have charts which reach over and beyond the two axes for the plotting of negative values. In still other charts, the true y-axis does not appear at all, the chart not showing no zero x-value whatsoever. This is generally a case of data in which zero itself is meaningless or arbitrary. A common example of it is historical data, such as the first illustration in the last paragraph, where the time values commenced with the year 1790—not with the year zero. The horizontal lines, and particularly the distances along the x-axis, are called abscissae. The vertical lines crossing the ends of the abscissae or points on the x-axis, are called ordin-

ates. All points having the same abscissae or values along the x-axis lie in the same ordinate or vertical line, and vice versa, all points having the same ordinates or values along the y-axis lie in the same horizontal line.

As we have seen, the curve chart requires data with two dimensions, the curve being plotted upon a field which has two dimensions. Along the horizontal dimension or x-axis you will find the values of the independent or x-variable, generally the stubs in your table of data. For each value of this variable, that is for each stub in your table, there is a corresponding value of the dependent or y-variable, namely the figure in the column beside the stub in your table of data. This y-value is plotted along the ordinate or vertical line from the given point on the x-axis (or abscissae, indicted by the stub) to the height upon the y-axis indicated by its value. Another way of expressing this is as follows: In the data for a curve-chart each figure to be plotted has two values, one being the value of the figure itself and the other being the value of its stub in the table. These two values of a figure describe the co-ordinates of the point by which the figure is plotted on the chart.

Not only can the point be plotted from the data showing its co-ordinates, but the process can be reversed and the co-ordinates of a point can be read from the plot or chart, merely by following the intersecting lines through the point to their respective axes. For it will be seen that every point on the paper has two co-ordinates, one of which is the abscissa or horizontal line passing through it, and the other of which is the ordinate or vertical line passing through it. The point itself is sometimes called the "intersect" of these two lines. As two perpendicular lines can intersect at one point and one point only, there can be only one point described by any two co-ordinates. We can therefore locate or identify any point by its co-ordinates and the co-ordinates of a point may be said to fix rigidly its position.

If our data tell us that a certain town has a population of 3,000 persons in the year 1910, we should plot this population by moving along the horizontal or x-axis, to the distance or abscissa of the year 1910, and then moving upward along the ordinate or vertical line through that point, to the height of the horizontal line (or abscissa) passing through the point of population, 3,000, in the y-axis. The dot, or point on the paper which would indicate this town, would be placed at the inter-

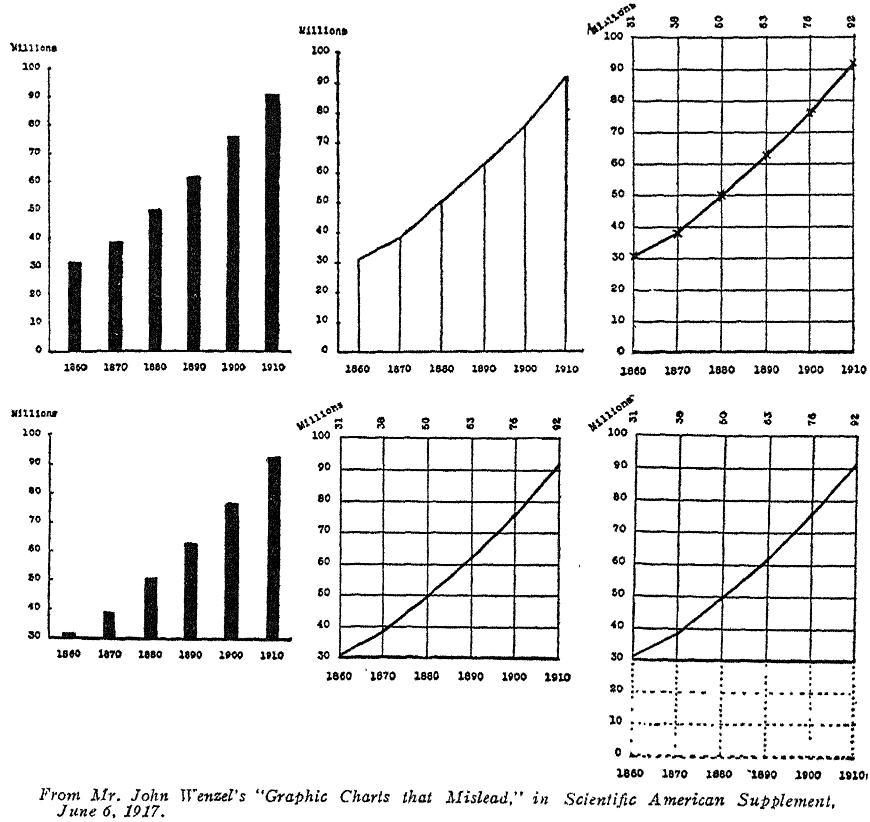
section of the ordinate for 1910 and the abscissa for 3,000, and the co-ordinates of that point would be "year 1910, population 3,000," that is "x, 1910, y, 3,000." And if we see the point of this town plotted upon a chart, we can read from its co-ordinates the information that in the year 1910 its population was 3,000, simply by following the two co-ordinates of the point out to the axes of the chart.

The distinction between the dependent and the independent variables is important. The independent variable is normally formed by the stubs in the tabulation and the dependent variable by the corresponding figures, that is, the figures in adjoining columns. The tabulation is more or less optional however, and for certain purposes stubs and data may be interchanged. The distinction between dependent and independent variables goes deeper, and finds its origin in the peculiar nature of the data. When the readings along any variable are made a basis of classification of data, then that variable is the independent one. In general, the dependent variable is that one whose values may be said to depend upon the values of the other variable. Such dependence need not take the form of a mathematical equation or explicit function, but is merely a matter of convenience in such matters as the classification and arrangement in the statistical table. Our chief concern is with the plotting of the data upon curves, and the important rule to be remembered is that the independent variable should be laid off on the x-axis and the dependent one on the y-axis. The rule is not without its exceptions, but these should always be founded upon special considerations and in the absence of such special reasons, the rule should be invariably followed.

Chapter XVI

FIELDS

Most of the good things in this world involve some sacrifice. Curves are no exception. In a curve the direct visible connection between the curve itself and the zero line, or x-axis, is sacrificed. As time goes on and you become more and more used to the curve chart, you will begin to think of its values

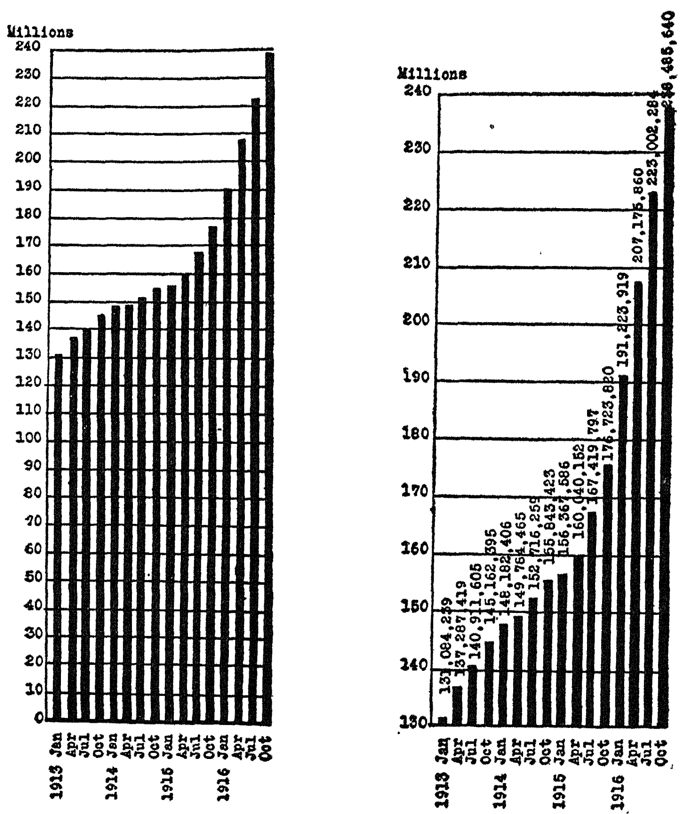


From Mr. John Wenzel's "Graphic Charts that Mislead," in Scientific American Supplement, June 6, 1917.

Fig. 138. The Amputated Chart is Deceptive.

as in some mysterious manner floating disembodied along the connecting line which forms the curve. You will be tempted to forget that the quantities rest very substantially upon the floor (base line, zero line, x-axis or whatever you want to call it), and that it is only their tops which reach the points plotted in the curve. And forgetting this, you will try to save space by omitting the zero line and lower part of the chart, and by showing only that small portion or band of the chart through which the plotted curve travels.

This practice of omitting the zero line is all too common, but it is not for that reason excusable. The amputated chart is a deceptive one, tempting the average reader to compare the heights of points on the curve from the false bottom of the amputated chart-field, rather than from the true zero line, far



From Mr. John Wenzel's "Graphic Charts that Mislead," in Scientific American Supplement, June 6, 1917.

Fig. 139. The Case Against Amputation is Clear.

below and invisible. A curve-chart without a zero line is in general no whit less of a printed lie, than a vertical bar-chart in which the lower part of the bars themselves are cut away. The representation of comparative sizes has been distorted and the fluctuations (changes in value) exaggerated. In a few more years, the principle that the zero line, when zero is real, must normally be shown in a curve will be universally accepted. Then the emphasis which now must be laid upon this principle, will not be needed. Indeed, the author plans in his fortieth edition of this work, to omit almost all reference to the rule. But, today, you will repeatedly find violations of the rule complacently propagating false impressions. And today the principle must be iterated, reiterated, and forever kept in mind.

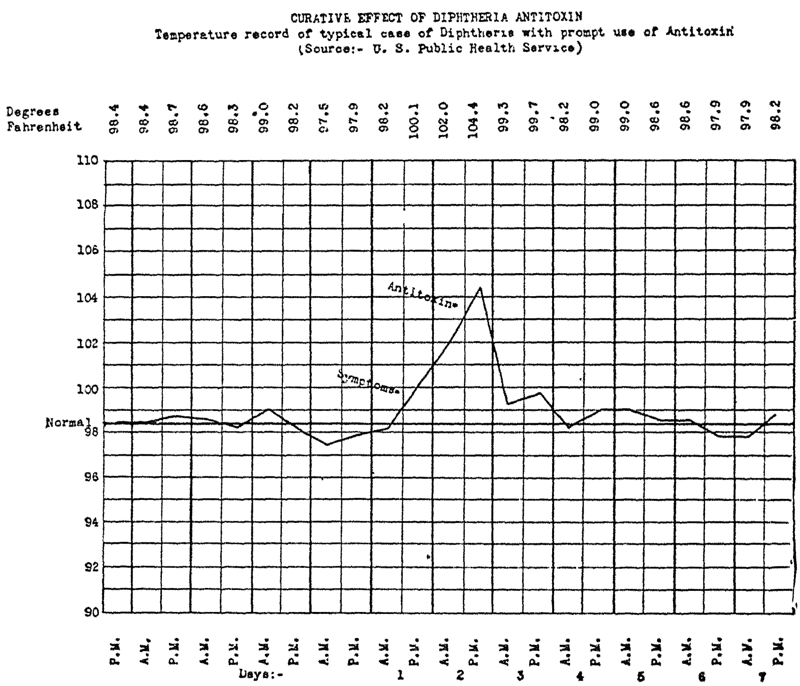


Fig. 140. When Zero is Arbitrary, it can be Omitted.

There is but one case when the omission of the zero-line on the y-axis or dependent variable, is justified. This is the case in which zero itself is an arbitrary value, and does not really mean a "nothing." As we have seen in the case of

x-values in historical data, the year O does not really signify zero years, but merely signifies an arbitrary point of time from which counting is begun. Science has many such arbitrary zeros; in the Fahrenheit scale of temperature, for instance, zero degrees is really an arbitrary point. Common sense will tell you when the zero is a starting point of the quantities measured by your data. And whenever zero is really such a lower limit, the rule that the zero line must be shown on a chart applies.

Sometimes, even with the best of intentions, rules must be violated and we must do the unjustifiable. The usual excuse for amputating a chart is that to show the zero line would require too much space, or would reduce the scale and make the fluctuations of the curve less noticeable. Sometimes you will feel the force of this argument very strongly. It is particularly frequent in charts of dividend, interest, and yield rates, where the fluctuations are in percentages and the base is understood by everyone to be 100%. The argument has greater force when the chart is intended chiefly for circulation among those in a profession who are already accustomed to think of the minor variations of percentages, and who would study the chart with interest only with regard to its time to time fluctuation-quantities, but would have no interest in its relative total quantities. Here it may be argued that the amputated chart would deceive no one and would be of greater service than if, at the cost of detail, it were made complete.

When this argument arises, it must be scrutinized with care and hostile scepticism. Often the argument will be found specious, resting more on the familiarity of the chart-maker himself with his data than upon the true attitude of those who will see the chart. In other words the maker of the chart, in the thoroughness of his own understanding of the data, forgets that others will be less familiar with it, and attributes to them his own skill and comprehension. It is easy, in this way, to be modest about one's own powers of understanding, but the modesty is costly when, as so often happens, the reader is loath to admit his inferiority, and merely lays the charts aside for study at some time "later on" which time, needless to say, never comes.

Only if it is quite certain that no misunderstanding will result, should the chart be amputated for the sake of saving space or exaggerating fluctuations. And even in such cases,

great care should be taken to make the amputation self-evident to the most casual reader of the chart, for it is precisely the man who has little time for study of the chart, who is most likely to be deceived by it. The best method of making the amputation of the chart obvious is to blot out with Chinese

WORKERS OUTPUT AND FATIGUE
Index Numbers of Hourly Output in Dexterous Handwork Operations
United States
(Source:- United States Public Health Service, Bulletin No. 106)

Commutator loom	85.2	97.2	97.2	100	95.5	94.3	93.2	86.6
Magneto taping	92.4	100	95.7	95.1	96.8	96.6	96.0	91.5
Roll coil	96.1	98.8	100	98.4	92.2	98.4	98.8	88.3
Rivet press	91.2	95.9	100	94.6	94.5	92.9	94.2	89.1
Average	91.2	98.0	98.2	97.0	94.8	95.6	95.6	88.9

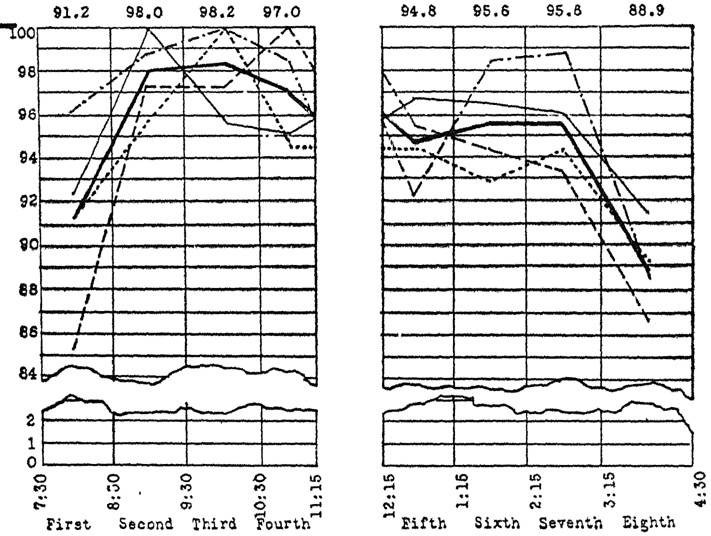


Fig. 141. The White Zone Warns the Reader.

white a small irregular zone across the lower part of the chart-field, or to erase the co-ordinates in this zone, and show the zero-line below this zone. The chart then has the appearance of being broken off between the zero-line and the curve, and anyone will see that it does not show full distances to the curve-points. An easier but less effective method is to make

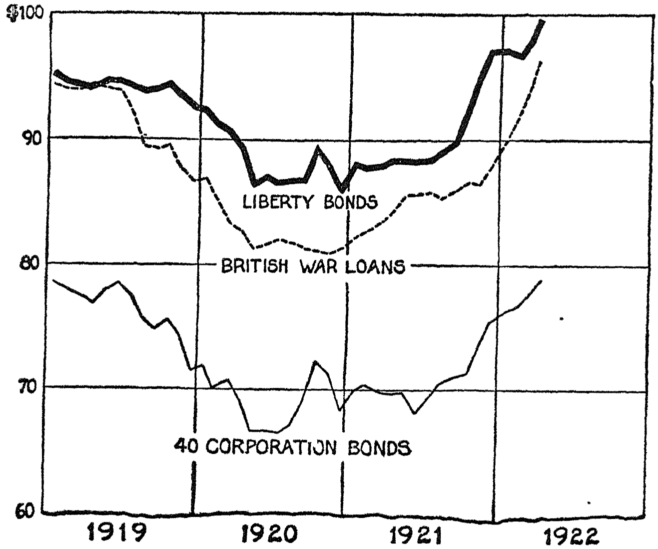


Fig. 142. The Uneven Base-line Indicates That It Is not the Real Base-line.

Average monthly prices of liberty and corporation bonds at New York and British War Loans at London.—*Permission of Mr. Carl Snyder.*

the lowest horizontal on the chart-field wavy or ragged, again with the purpose of indicating that a part of the chart has been broken off.[1] In either case the final object is to remind

Harvard Bureau of Economic Research.

Fig. 143. A Wavy Base-line is a Shorthand Warning.

Adjusted Index of the Volume of Manufacture (100 = Normal).

[1] The use of rounded or dotted base-lines to indicate abbreviation, which is sometimes advocated, seems ill-advised, since the method is not self-explanatory and hence defeats its own purpose. The object is to flash to the casual glance the abbreviated condition of the chart, and any symbolism which must be technically understood is of no more value than the scale-figures themselves for this purpose.

the reader than the true zero-line, from which he should measure the quantities shown by the curve, lies far below the visible portion of the chart.

Another principle which will quickly appeal to your common sense, is the rule that when zero is real, the zero-line should be extra heavy to make it prominent. Remember that it takes the place of the floor or lower end of the bars in the bar-chart. It should stand out, therefore, in such a way that the reader can easily grasp its significance and compare with it the heights of the points on the curve. The rule is particularly important in cases where the chart extends down below the zero line into the negative side in order to show negative and positive values. On the same principle the 100% line, when it occurs in a chart, should be similarly heavy as it also may be

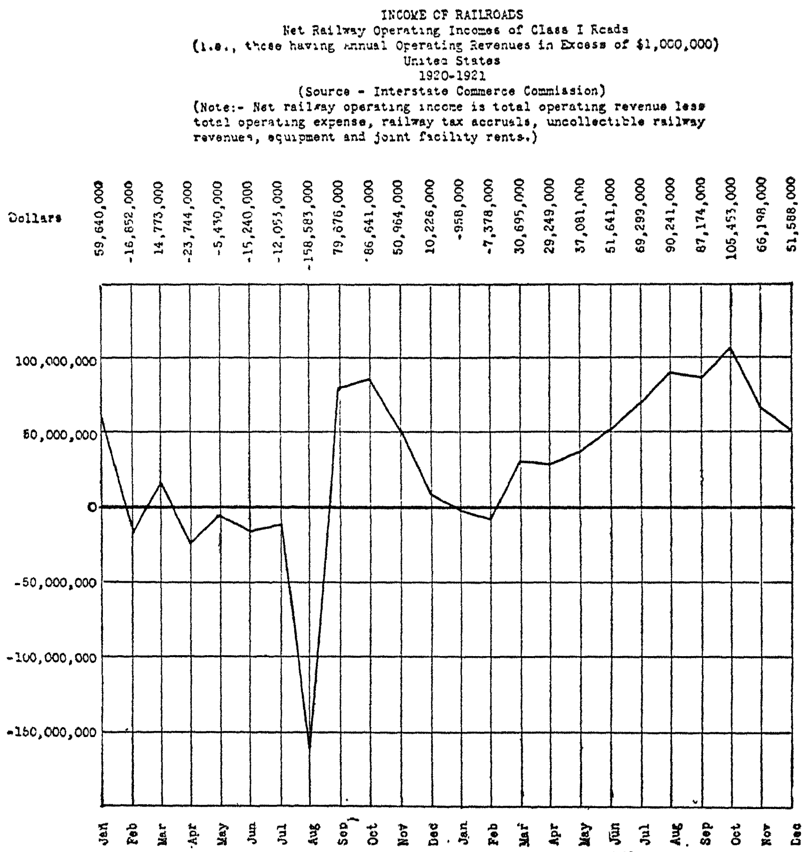


Fig. 144. The Zero-line Should Always be Heavy.

considered a base for zero points, being the point of zero loss or gain. In fact, the rule may be extended to all cases of lines showing significant constant values, and the zero line should not be heavy, unless it has a special significance. In charts showing temperature in Fahrenheit degrees, for example, since the zero point is merely an arbitrary value like any other number of degrees, it would be more sensible to emphasize the freezing and boiling point lines. Common sense must be relied on to determine the lines which can be usefully emphasized.

The ordinary curve-chart has three different types of figures which must be attached to it. These figures are: first, the scale figures for the x-axis, showing the values assigned to the vertical lines; second, the scale figures for the y-axis, showing the values which have been assigned to the horizontal lines; and third, the data figures, showing the values represented by the various plotted points of the curve. There is considerable confusion and difference of practice in the positioning of these three sets of figures. By going back to the first principles, however, and recognizing that a chart is merely a fragment of the co-ordinate system of measurement, we can easily find the logical and natural places for these figures, and it so happens that the positions which are the most logical have proved in practise the soundest and most useful ones.

The scale-figures along the horizontal or x-axis are really the values of the independent variable. In your table of data they form the stubs or items. As you read in the last chapter, this independent variable belongs on the x-axis; it should never be placed along the y-axis, as is sometimes erroneously done. The proper place for the scale showing these figures or values of the independent variable, is at the bottom of the chart, each figure or value being immediately beneath the lower end of the vertical line to which it has been assigned. Do not, merely for the sake of ornamentation or decoration, place these values at the top of the chart also. Do not box them in, each with a little square or circle. Do not make the printing unnecessarily large. Do nothing more than is necessary for simple clear results. These precepts will save you a great deal of time in the preparation of your charts and will save your reader much trouble in its reading. It is enough to place the figures once for all at the bottom of the chart, forming a scale along its entire base.

The figures which are assigned to the various horizontal lines should be placed in a column immediately beside the ends of these lines. They form the scale for the dependent variable or y-axis. In the case of isolated charts, that is charts which will appear singly and alone, this vertical scale is often placed at both sides of the charts so that the reader can read the values of the horizontal line at either side. For isolated charts, there is no particular objection to this practise, though it may be a work of super-erogation. But, in the majority of cases your charts appear in groups, and often on separate sheets which the reader will wish to place side by side for the purpose of comparison. Then certainly the two vertical scales would be a nuisance; one is sufficient, and it should be placed at the left hand side. Indeed, it would be best to make this rule universal, namely, that the vertical scale should appear

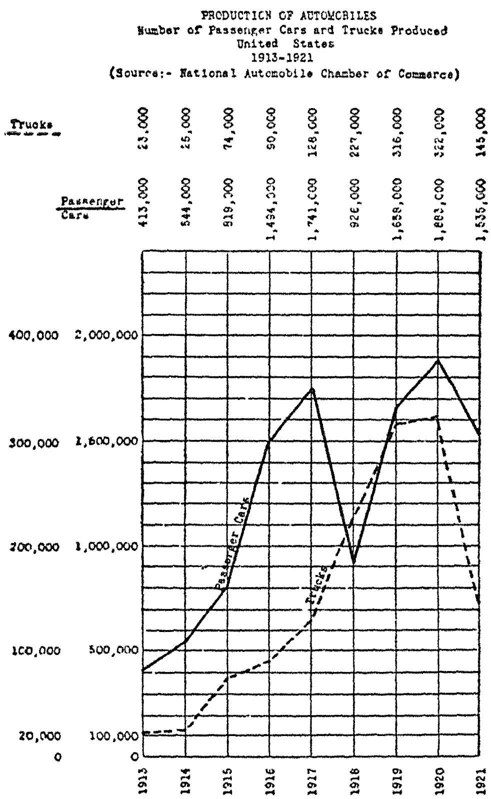


Fig. 145. The Sound Position for Two Vertical Scales.

once only, and then at the left hand side of each chart. The use of two scales, one on each side of the chart, is desirable only for more popular results.

When several curves are shown upon the same chart, it is often desirable to use different scales for them. That is, the same horizontal lines may be given two or even more different values for different curves. But even in these cases, it is better to place both scales, once and for all, at the left hand side. The practise of placing one of these scales at the right hand side, and another at the left hand side, has little to recommend it. Theoretically, at least, the léft hand end of your chart is normally the y-axis itself, and the scale or scales should logically be attached immediately thereto. In practice this logical position is justified.

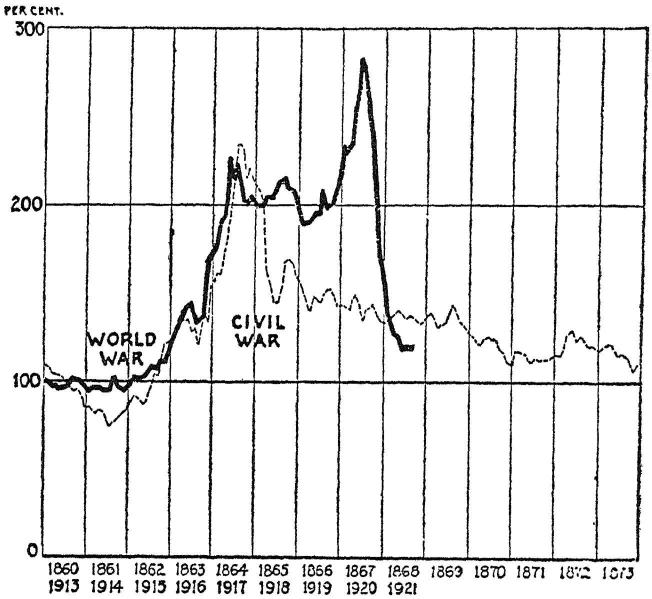


Fig. 146. An Interesting Comparison of Different Periods.

Here it is not the y-axis, but the x-axis, which has two scales.

Monthly price index of 14 basic commodities during two war periods. Pre-war year in each case is taken as the base of 100%. Prices of the same commodities are included for each period.—*Permission of Mr. Carl Snyder.*

We come then to the question of the third set of figures, namely, to the data itself, which the curve represents. As we have said before, this data should always be presented with

the chart.[2] It should not be omitted entirely or separated and printed in an appended table. That rule is one of the most important in chart work, and he who violates it fails to afford his reader with convincing proof of the accuracy of his chart. We must find a way to insert the data in the chart, and typographical difficulties or inconveniences must not be allowed to deter us.

As a matter of fact, the place for the data is obvious. It should be placed at the top of the chart, each figure immediately above the point by which it is represented on the curve. The reader can then glance down the ordinate or vertical line through any particular point and find the stub or value of the independent variable, and he can glance up the same ordinate or vertical line, and find the value of the curve at that point, that is, the exact value of the dependent variable. Needless to say, he could have found this latter, the dependent variable value, with approximate precision, by careful study of the horizontal line through this point and its intersection with the y-axis.

In entering figures on the chart for the scales of the independent variable and for the data itself, we come to the same typographical difficulties which we met in the pipe-organ or vertical bar-chart. We do not often find sufficient room to print or write these figures horizontally on the page. Even if the chart is so large that we can, by fine printing, crowd the figures together horizontally, we will generally find the results unsatisfactory. In the first place the figures tend to run into each other, and in the second place they lie across, rather than in line with the ordinates or vertical lines to which they are attached. If we attempt to box the figures in with squares, circles, or diamonds, we merely add to the confusion of the chart and detract from its simplicity.

---

[2] It is obviously the chart from which data has been omitted, which has led Professor Secrist to say:

> "Tabulation of classification precedes; the use of diagrams follows. The former generally serves to clarify the meaning of data; the latter frequently to obscure it . . . Diagrams alone are more likely to serve as bases for conclusions arrived at without study and to foster a disregard for the details from which diagrams are drawn . . . Diagrammatic illustrations can never replace data themselves, no matter how accurately they tell the truth or how illuminating they are. They are at best statistical aids and should be so viewed by those who use and study them. A well-drawn and cleverly executed diagram is never a guarantee of the value of the statistical facts which it illustrates."—Secrist, Horace, *An Introduction to Statistical Methods*, The Macmillan Company, New York, 1917, pp. 159, 161.

The sound principle, therefore, and one which will be found, after a little practise, eminently satisfactory, is to enter the data-figures and all except the simplest x-scale figures, vertically, that is, by writing on edge. The reader has little difficulty in turning the page about to read these figures when he wishes. There is, therefore, no great disadvantage to this method. The figures lie clearly along the lines of the ordinates to which they belong, so that there is no doubt or confusion in finding the figure for any particular point on the curve. You will notice, moreover, that the figures for the independent variable and the figures for the data arrange themselves in the familiar form on your original tabulation, the only difference being that a chart has been inserted sideways between the stubs and data. And this is logically sound, because the curve is merely a modified form of the pipe-organ chart, and the pipe-organ itself is merely a bar-chart placed on its side. Even the column headings are retained in the curve chart in their same relative positions to each column of figures.

You will find it useful to keep this relation between the curve-chart, the pipe-organ or vertical bar-chart, and the bar-chart proper or horizontal bar-chart, always in mind. Particularly so, when you have several columns of data to be shown by several curves upon the same chart, for in this case it is important to retain the column headings at the top of each column of figures in the data. These column headings will then be to the left of the chart itself and the only difference will be that they can be written on horizontal rather than vertical lines, so that they can be read easily while the curve-chart is in its normal position, though the data is on edge. You will, however, find it useful to re-arrange the order of the columns of data so that the position of each corresponds roughly to the position of its particular curve, the data for the uppermost curve being at the top, and the data for the lowest curve being at the bottom of the series of data on the chart. In order to distinguish two curves on the same chart, which may or not cross each other, you will probably use different colors for these curves. In that case, it is useful to observe a similar color distinction in the printing or typing of the data columns, each column being printed in the color in which its particular curve is plotted. You will also find it useful to place a small sample section of this curve immediately beside or underneath the column heading for the data to which it is

attached, thus forming a sort of key to the curves used on the chart.

For those who use typewriters (and in all large offices, it is well to use typewriting exclusively for the lettering and figure-writing on charts) the arrangement above described for the positioning of figures and data, will be found extremely convenient. It gives the typist no more trouble than the preparation of an ordinary table or tabulation of figures. Placing the chart sideways in the typewriter, she types in at the left hand edge of the chart-field the stubs or items of the table, and at the right hand end of the chart, the figures of the various columns. If the ordinates of the chart have been arranged at the precise typewriter distances of $\frac{1}{6}$, $\frac{1}{3}$, or $\frac{1}{2}$ inch apart, she has no further adjustment of the paper to make in the typewriting machine. In line after line down the page, she merely reproduces the table of the original data leaving a wide gap between stub and columns of data—a gap which is filled up by the field of the chart. The chart can then be plotted in upon this field after the data has been typed.

The whole process of making a curve chart takes no more time than that of making a bar-chart, and in fact, very little more than that of making a plain mathematical table. The only instructions the typist must have are the data (which can be in the form of the original table) and clear orders as to (1) which columns must be copied on the chart, (2) the order in which they must be placed, and (3) the color, if two colors are used, in which they must be typed. The only instructions which the draftsman needs are then contained in the form itself on which he is to draw, his instructions being the data as already typed on the chart. If the vertical-scale figures for the y-axis of the chart have also been typed in, his instructions are complete. But unless the chart belongs to a standardized set in which the scale has been fixed, the draftsman will probably determine upon his y-axis scale after a study of the data itself. In this case, he enters the scale in hand-lettering and proceeds with the plotting. If this scale also is to be typewritten, however, he would enter the scale-figures only in pencil and the typist would enter the figures permanently last of all. The considerations affecting the choice of scale will be found in the following chapter.

Chapter XVII

SCALES

Technicians are fond of describing a scale in puzzling and abstruse language. Yet, as often happens, the thing itself is so simple that a child can understand it. It is usually defined as a ratio—the ratio between actual distances in the space charted and equivalent distances on the chart. This ratio of reduction or enlargement is important to engineers but not to the maker of mathematical charts. We therefore use the word scale for the calibrations measuring distances on the chart. To linear distances, both horizontal and vertical, we assign arbitrary values and the figures which tell us these assigned values form the scale.

Curve charts take up two dimensions on the paper, that is, they have both a vertical and a horizontal axis, and therefore require two scales. These two scales may or may not be alike. When they are alike, we have what might be called a normal projection. Imagine a simple chart, the field of which is square and the two scales of which are alike. Draw a straight line from the point of origin or lower left hand corner of the chart-

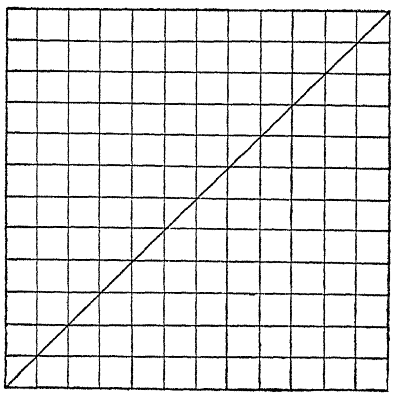


Fig. 147.

field at an angle of 45° to the horizontal and extend it diagonally across the field to the upper right hand corner. This line passes through all points having equal co-ordinates, that is equal values along both axes. Now let us see what happens to this line when one or the other scale of the chart is changed.

Suppose we shorten the vertical scale to half of its distance. Relatively speaking, this is the same thing as doubling the horizontal scale. (By half or double the length of the scale, we mean assigning the measurement values to distances half or double as great.) Now the result of this will be very noticeable upon the slope of the straight diagonal line passing

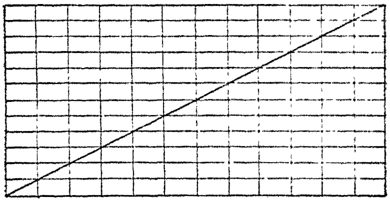


Fig. 148.

through points having the same values as before, for the line will rise only half as much as before. If our line were a curve wiggling across the chart, its wiggles would be half way

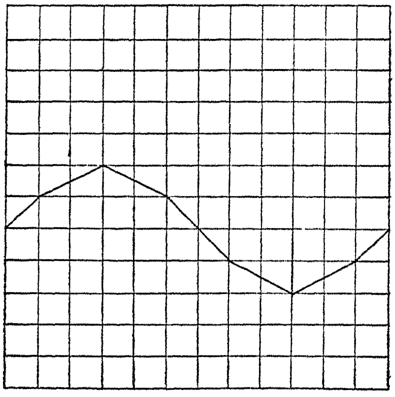


Fig. 149.

flattened out, giving us the impression of much less fluctuation than formerly. But as a matter of fact it is exactly the same curve as before, only its field has been changed so as to diminish the vertical oscillations or wiggles.

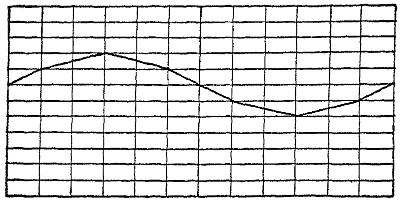


Fig. 150.

On the other hand, suppose that we increase the vertical scale to twice the length of the horizontal scale. This is the same thing, relatively, as reducing the horizontal scale to half-

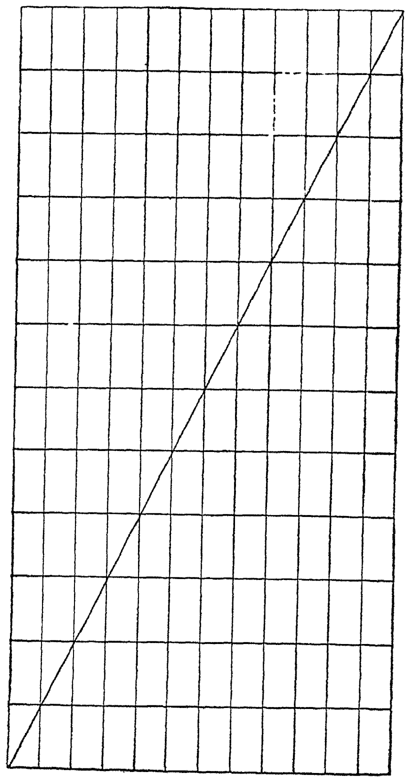


Fig. 151.

size. Now see what happens to the diagonal line. Its slope becomes far steeper than originally, as it must climb to twice the height in the same horizontal distance. If that line had been a curve, snaking its way across the paper, its wiggles

would have been twice as great as formerly. It would have given us the impression of a very unsteady and changeable

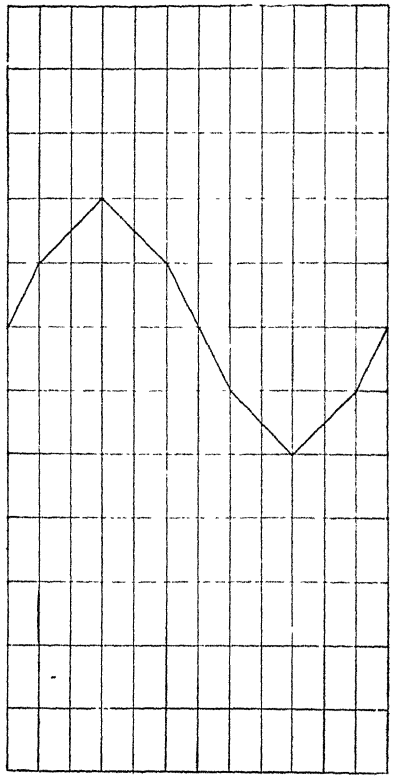


Fig. 152.

proposition indeed. First way up, and then way down. Very hard to tell just where it is going to go next. Unstable, unreliable, fickle—these are the conclusions we should have formed of the items that were charted, and yet those items are precisely the same as appeared on the second chart above described where their movements appeared to be very even and regular.

In short, the scales on which a curve is drawn can affect very much our impressions of the data by magnifying or minimizing the apparent movements of the curve itself. Of course, this does not mean that the relative height from the base-line of the various points on the curve have been altered. If you have been careful to show the base-line always, the base-line

itself will approach nearer to the curve as the vertical scale is reduced and the wiggles are flattened out, and will recede

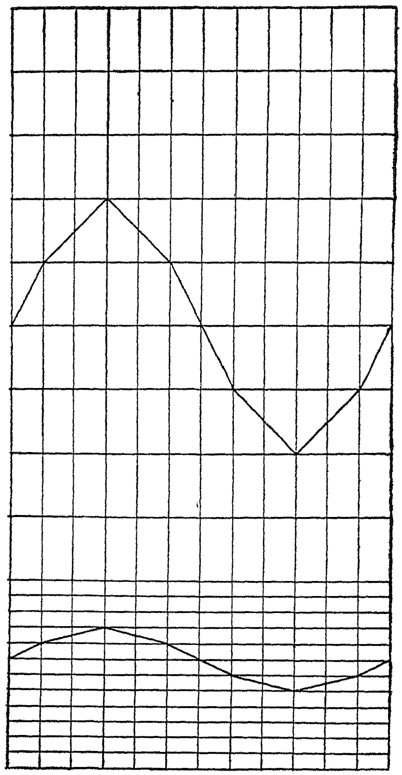


Fig. 153.

farther from the curve as the vertical scale is enlarged and the wiggles are exaggerated. But it means that the oscillation or fluctuation of the curve will have been made to appear more violent or milder according as either of the scales is changed. And it therefore behooves us to give serious thought to the matter of scales before we determine upon them finally for any particular chart. As a matter of fact, we may have to try out several combinations of scales before we find one which gives just the right amount of emphasis to curve fluctuations to suit us.

Now where our chart size is unlimited, and we are free to extend the scale and field in either direction as far as we wish, this rule of try, try, try again might be perfectly feasible.

Perhaps, in that case, we would generally come back to the normal projection or combination of two similar scales. If, as often happens, the scales measure different and incomparable (technically "incommensurable") quantities, such as years on one axis and dollars of sales on the other, or length in inches on one axis and weight in pounds on the other, we could change these to percentages (each of its own total or maximum), and consider the percentages commensurable.

But generally, the space available for a chart is limited. If it is to appear in a book, the size of the book-page must be conformed to. If it is one of a set of charts, a uniform chart-size increases the attractiveness, if not also the simplicity, both of the set and of the individual chart. Even if the chart is to appear entirely alone, there is much benefit in avoiding unhandy sizes. Moreover, worrying through a succession of trials consumes time and energy, a needless waste if it is true that we can determine beforehand merely from the data itself what will be a satisfactory combination of scales.

Let us consider the horizontal scale first. In the previous chapter we have already found certain considerations which will affect the arrangement of the figures for this scale. These figures will be placed immediately below the base-line or bottom of the chart. Normally, they will be written on edge, upward, or typewritten after the paper has been fed into the typewriting machine sideways. The ordinates or vertical lines to which the figures belong will then, if extended down the chart, pass through the figures, cutting across the middle of each digit. In the same way, above the top of the chart the data-figures will be placed, each on line with its own scale-figure and plotting point and each so placed as to be similarly cut by the extension of its ordinate.

It takes no brains to see, therefore, that the horizontal scale must be large enough to permit entering the figures, no matter how condensed, of the data. As a general rule, typewriter intervals, which are in picas or sixths of an inch, are about as small as your horizontal unit-distances should be.[1] And if you have a short series of data, you can double or treble this distance without expanding your chart too much. In fact, the curve is more easily read when the horizontal

[1] All typewriters can be especially equipped, at a slight extra cost, with any desired interlinear distance and there is one machine, the Hammond, frequently used in academic work, which has intervals of one-ninth instead of one-sixth of an inch.

units are about typewriter double-spacing distance apart, that is, three to the inch.

We are assuming here that your finished chart is to occupy a sheet of paper about standard letter-size, 8½ by 11 inches. If larger sheets are to be used, you will modify all dimensions accordingly. Where charts are to be exhibited in a large room to a large audience, they must be many times larger and all

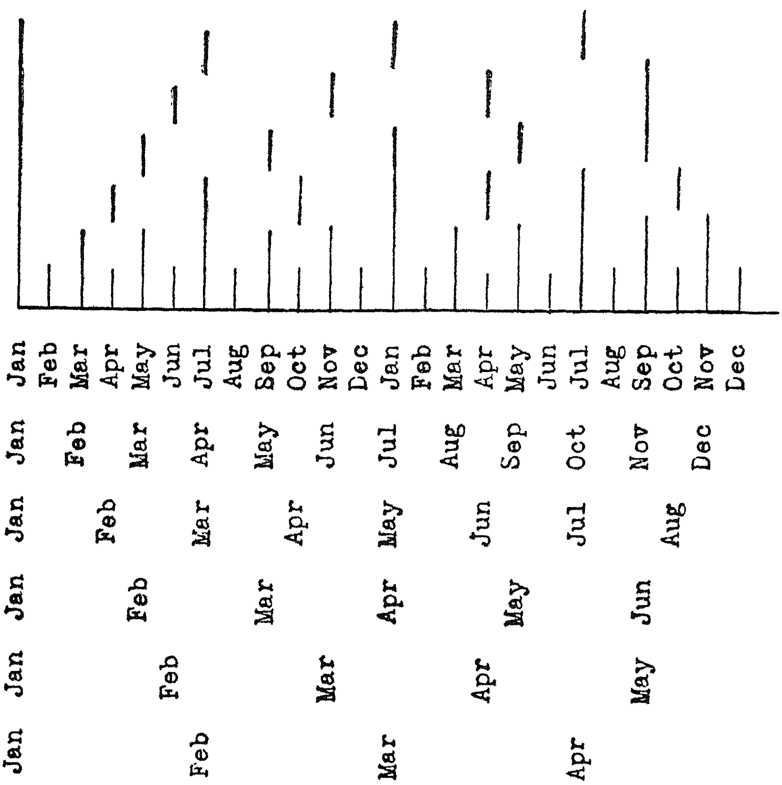


Fig. 154. Examples of Convenient Horizontal Scales.
Facsimile Typewriting.

lines and lettering correspondingly heavier. The most excellent chart in the world is virtually useless to the man who cannot see it, and you must not forget the distance from which the chart is to be viewed. For ordinary study, however, as well as for convenience in handling and in filing, the 8½ by 11 basis is satisfactory. It can always be enlarged by photo-

stats, or photographed on lantern slides for very large projection.

The range of this horizontal scale then depends largely on the number of items in the series, to be plotted. A series which contains more than thirty items had best be cut up into

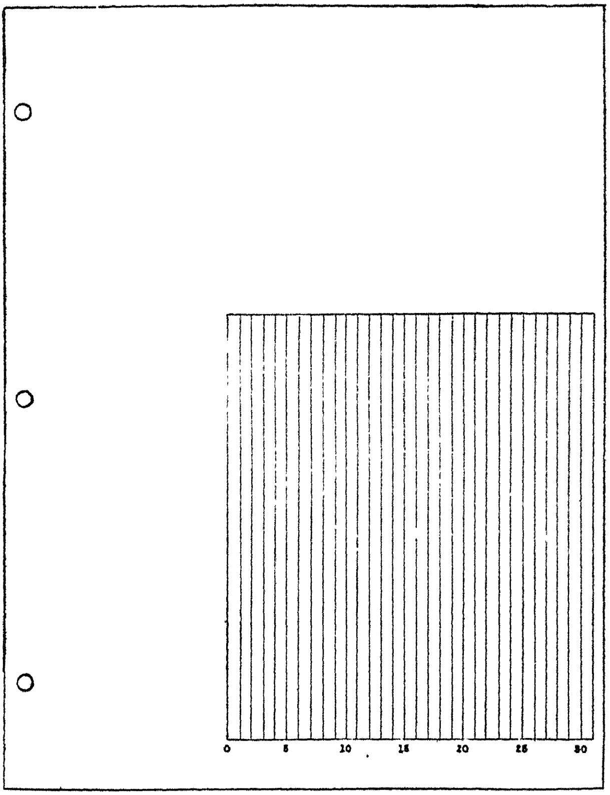


Fig. 155. Showing One Month by Days on Letter-size Paper.
Single-spaced for typewriting data.

two charts, each one of which will run across the shorter distance of sheets of the 8½ by 11 paper. You can generally do this by breaking up the series into convenient segments. Thus if the data is monthly, break it up into years and present a year on a page. If the data is annual, break it up into ten or twenty year groups. Where it seems inadvisable to break up the series into parts this way, double width sheets can be

used, either folding up into regular size or not, as desired. If you wish to run the chart along the long distance of the 8½ by 11 paper, the space for attaching data above the chart will be much restricted, but if the data is limited to one or two columns, this is no disadvantage, and you can get as many as

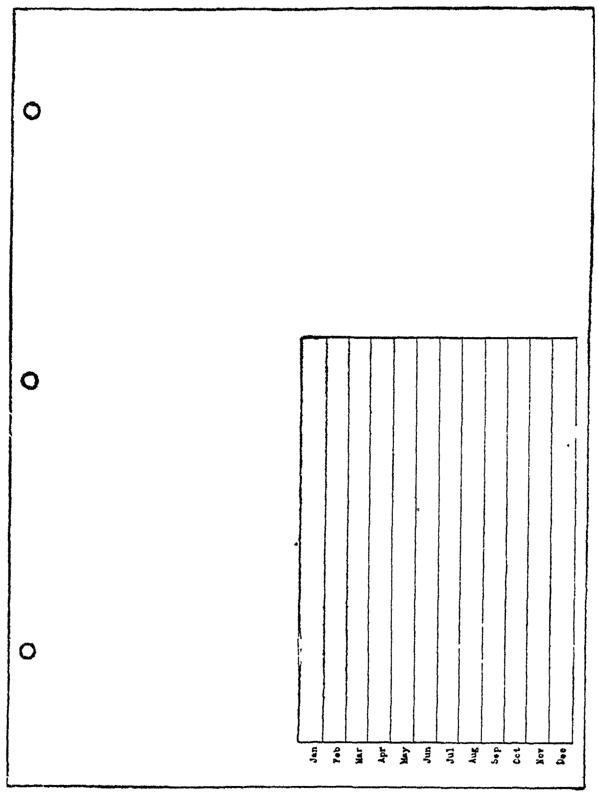


Fig. 156. One Year by Months on Letter-size Paper.
Double-spaced typewriting.

fifty-two items on a page, thus enabling you to show a year by weeks.

When you can do so, it is always well to make the horizontal distances one-third inch each, or double typewriter-spaced. In this case, you cannot count on more than a dozen or fifteen units crosswise on the paper and twenty-five lengthwise. The advantage of the wider spacing, as has been said,

lies in the greater ease with which it is read, neither its curve oscillations nor its data figures being so confusingly close together as when smaller spacings are used. Moreover, the chart with fewer items on it will generally be more closely studied by the reader than one with a great mass of detail. In fact it sometimes pays to omit minor details in the data and make the items fewer and more important, in order to reduce a great amount of detail to a simple series. Thus the daily stock quotations would require a very large chart for their presentation during a year, while the weekly and sometimes the monthly average quotations will be just as significant, and far simpler, to the reader.

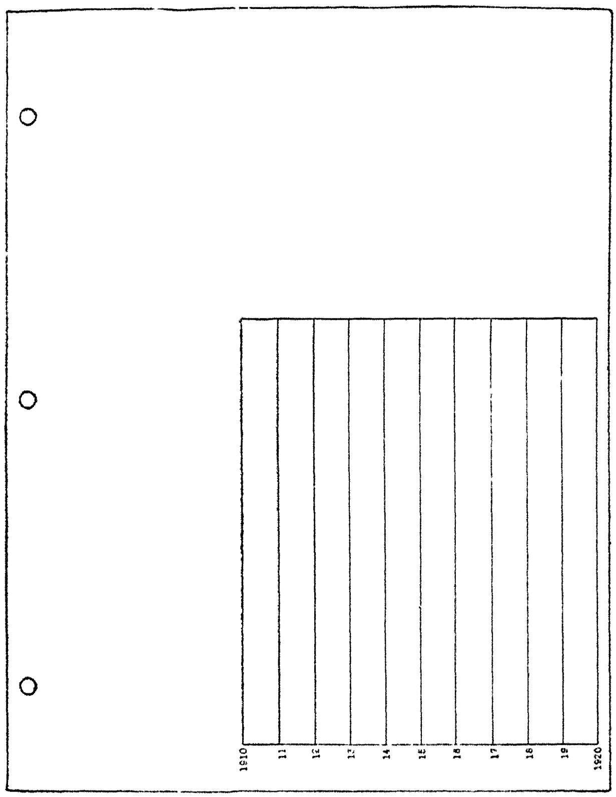


Fig. 157. One Decade by Years.
Triple-spaced typewriting.

Now as to the vertical scale.[2] The first general rule is that the highest plotted points on the curve should ordinarily reach about two-thirds of the way up the field of the chart. This gives the best results, because the top of the chart neither

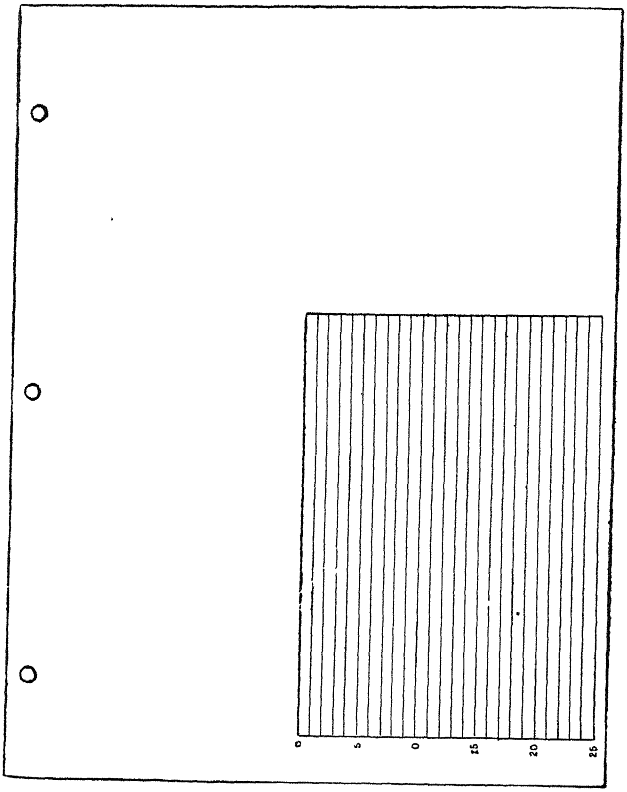


Fig. 158. One Quarter-century by Years

Single-spaced typewriting.

crowds the curve too closely nor does the space above the curve seem to the reader unnecessarily large. If the top of the chart is too close to the curve at any point, the reader may be

[2] It is to be understood in the following discussion that what applies to the positioning of a single scale for a single curve applies also to two or more scales for two or more curves when these are shown on one chart. Unless there is special reason for having one curve below the other, or for using a common scale for both curves, the second curve may have its own scale (lettered on the chart beside the first scale) specially positioned, like the first scale, to bring the second curve to similar heights upon the chart.

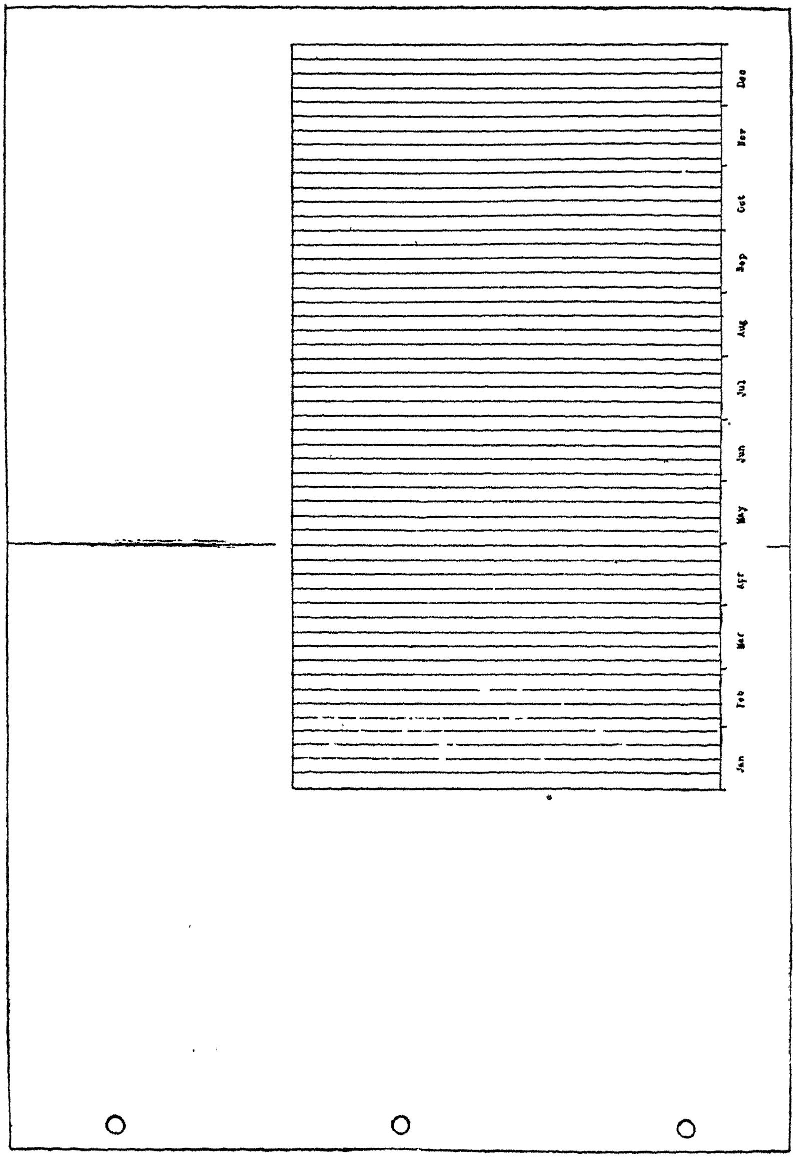


Fig. 159. One Year by Weeks on Double Letter-size Paper.
Single-spaced typewriting. The long line shows the fold.

led to measure with his eye distances on the chart from the curve to the top line, instead of from the bottom line.

When a single long series of items is to be carried through many charts, one after the other, forming a set in which the individual charts show only parts of the series of data, it is important to have the vertical (as well as the horizontal) scales uniform throughout. The uniform scales are necessary that

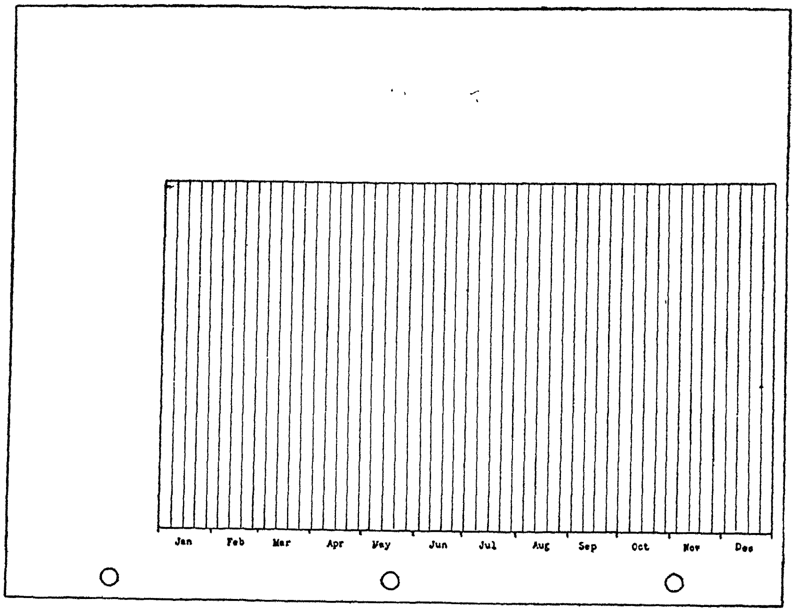


Fig. 160. One Year by Weeks on Letter-size Paper.
Single-spaced typewriting.

the charts may be individually compared, or "fanned out" into one long series of continuous charts. And if the scale be such as to place the highest point in the whole series three-quarters of the way up the page, in one chart, there may be other charts in the series, in which the curve will hardly leave the zero, or base-line. This cannot be helped without enlarging the scale for these smaller parts, and so destroying the comparability of the charts. There is no help for the low charts in this case, nor is help really desirable, since the lowness of the curve at certain points is the significant fact to be shown.

The size of the vertical scale depends therefore upon the amount of the largest figure in the date. We must glance

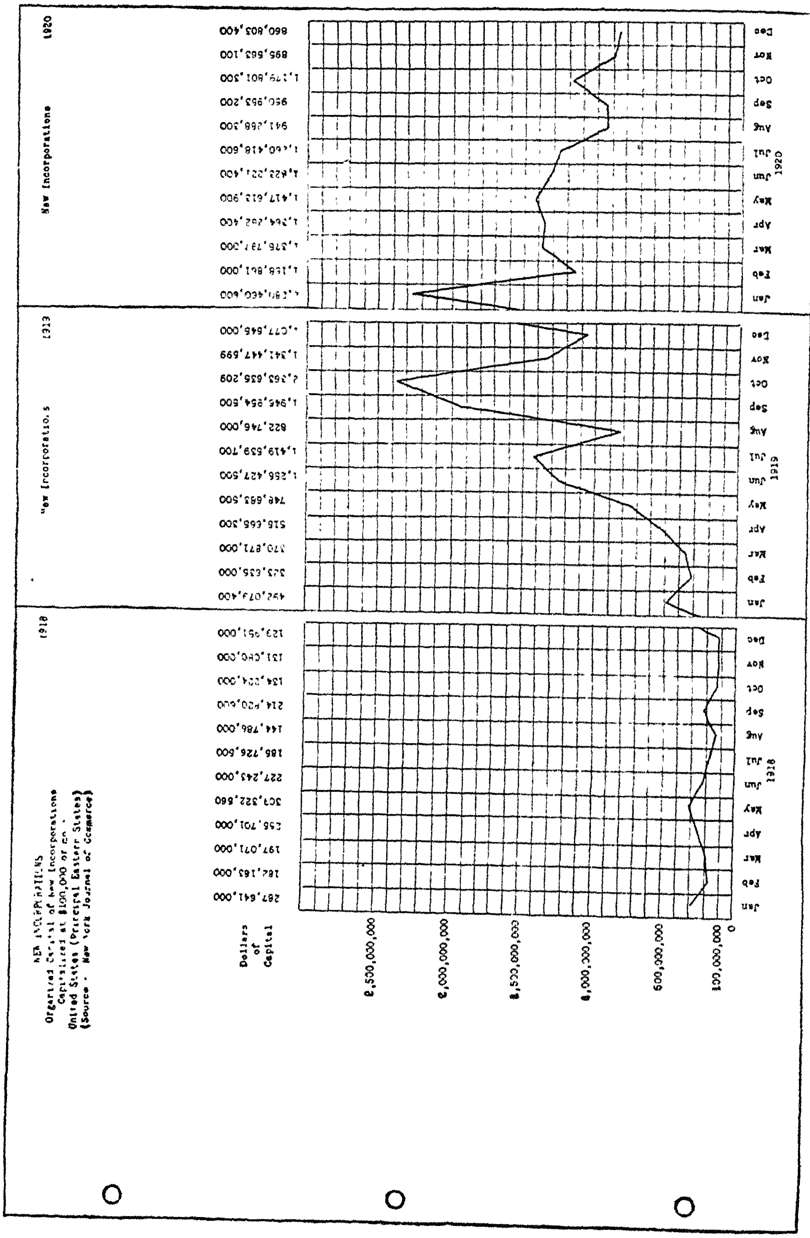


Fig. 161. Several Charts Overlaid Appear as One.

through the columns of data to be charted and observe the highest quantity in the series. Of course if this is a freak quantity, we can disregard it and select the next highest quantity (leaving the highest one to extend clear out of the chart if it will). Having determined on what we shall consider the high point or "peak" of the data, let us substitute for this a round figure, which we shall position about two-thirds or three-quarters of the way up from the bottom of the chart.

FIRE LOSSES
United States
1875-1920
(Source:- Journal of Commerce)

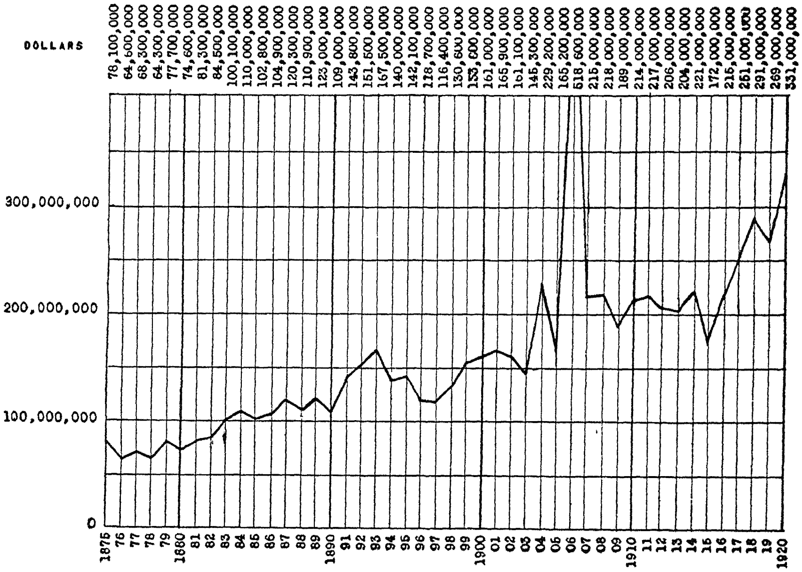


Fig. 162. The Freak Peak Need Not be Accommodated.

Thus if the high-point or peak is $8,370,000, let us take $8,000,000 as the scale-making figure. Now if this peak is approached by several others, that is, is no unusual value, let us make it slightly lower down on the chart, but if it is unusual, and the general level is nearer six or four million, let us make it slightly higher up. Anything between a third and a quarter of the distance below the top of the chart is sufficient. Assuming that the series contains several figures near eight million, we will select the slightly lower position, and place eight million two-

thirds of the way from the bottom of the chart to its top. In other words the entire vertical distance will be divided into twelfths, each representing one million dollars. In this way our vertical scale has been determined.

Of course, the size of the chart itself has not yet been settled. Its width we disposed of under the head of horizontal scales. But so far we have not settled its total height. We have only decided the number of parts into which that total

Jan	5,902,000
Feb	5,640,000
Mar	6,413,000
Apr	6,494,000
May	6,309,000
Jun	6,186,000
Jul	6,329,000
Aug	6,251,000
Sep	6,350,000
Oct	6,819,000
Nov	6,147,000
Dec	8,370,000

United Cigar Stores Co. Sales
1921
(Dollars)
(Source:- Survey of Current Business)

Fig. 163.

height will be divided. But the total height need give us little trouble. For a drawing on ordinary letter-size paper, the chart field, that is, the co-ordinate rulings, should not cover much more than half the height of the sheet of paper. There will be some space needed at the bottom for the horizontal scale figures, and considerable space should be left at the top for the data and for the title to the chart. So on a sheet of paper 11 inches high, the chart can best be made about six inches high. And in the example we have just considered, where this total height of six inches will be divided into twelve parts, each representing a million dollars, it is easy to see that the ordinates or horizontal lines should be spaced half an inch apart. The scale ratio is $\frac{1}{2}$ inch to $1,000,000.

There are many devices for dividing a length into desired divisions. The case of ten, fifteen, or eighteen or more divisions is no more difficult in a six-inch space than that of twelve divisions. Both engineers' and architects' rules divide the inch into various useful numbers of parts and when we desire an odd or fractional number of parts per inch, we can

United Cigar Stores Co. Sales
1921
(Dollars)
(Source-- Survey of Current Business)

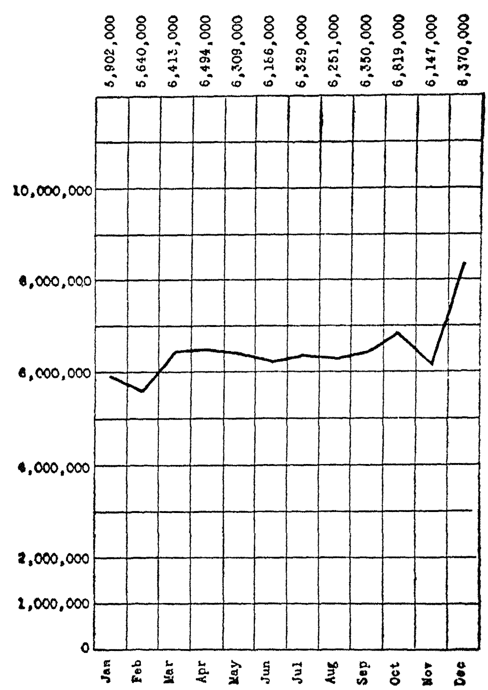


Fig. 164.

easily get them by drawing parallels from the corresponding divisions on a regular scale laid off so as to form a triangle with the desired scale and the last parallel. However, as we had considerable margin of choice in deciding the number of dividions, we can always find round figures which will work easily in the given chart field.

The fact is that a standard "field" about 4 inches wide and 6 inches high, has already been adopted by a great many chart-

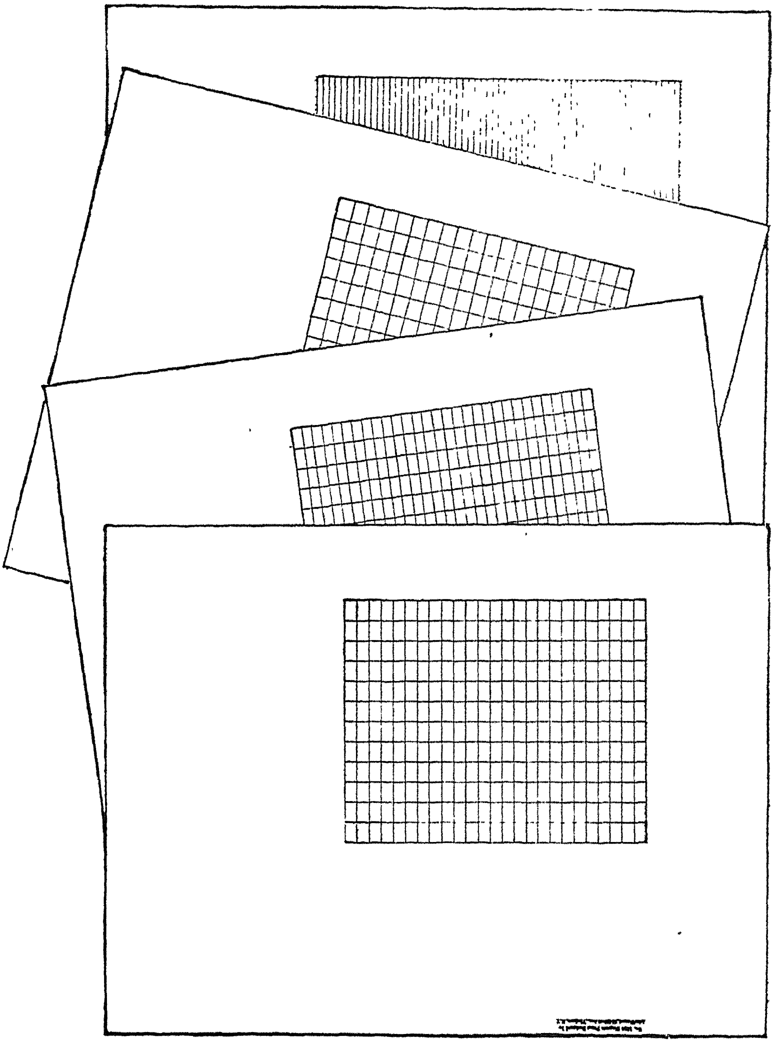


Fig. 165. Commercial Forms Available.
Four Useful Charting Papers published by Mr. John Wenzel, Yonkers, N. Y. The first three are of the type described in the text to fit the 10 or 50, 30 or 60, and 20 or 40 sides of the Engineer's rule. The fourth is specially adapted to percentage data, requiring a scale from 0—100%.

makers, and by some publishers of chart-paper. This standard p·epared chart-form is very useful. It is printed low upon regular letter-size 8½-by-11-inch sheet, leaving the necessary space below the chart for horizontal scale figures and a great deal of space above the chart for data and title. It is generally printed with the horizontal rulings only, so that any desired number of vertical rulings can be drawn in to suit your hori-

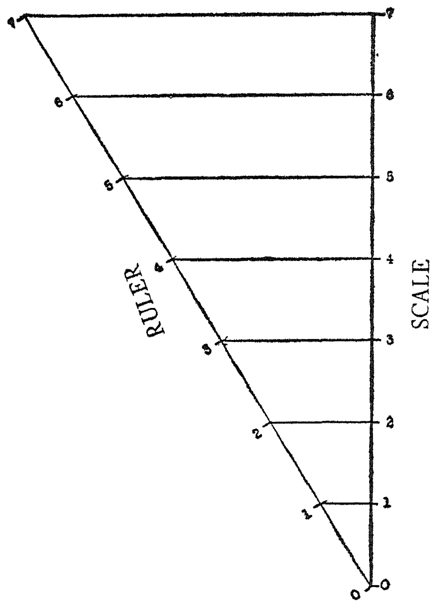
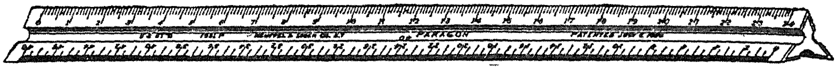


Fig. 166. To Obtain a Scale Smaller Than Those Given by the Ruler.

zontal scale. And the horizontal rulings are printed without any scale-figures for the *y*-axis, so that you can adopt whatever vertical scale you please. Three different rulings are made, in which the interval is either a fourth, a fifth, or a sixth of an inch. These three forms are sufficiently different to enable you to place a point almost anywhere you wish on the field, merely by selecting the right ruling and attaching to it the proper calibrations or scale-figures.[3]

[3] A single chart-form, which has intervals of one inch up the paper between horizontals, can be conveniently used in place of the three, when only a few charts are to be made. It is a master-form which can, if desired, easily be converted into any of the three by ruling in the proper number of intermediate horizontals.

Moreover, if you are doing much plotting, it will help you greatly to use what is called an "engineer's scale" ruler, which can be obtained with the inch divided into fourths, fifths, and



Courtesy of Keuffel & Esser, N. Y.

Fig. 167. Engineers' Triangular Rule.

sixths, coinciding with the three types of prepared chart-paper, and into other even fractions, namely halves, thirds, and tenths of an inch. With these six scales suitable for use with the three types of ruled paper, you can conveniently draw up any curve in any position on the field. And if you have not access to the specially ruled paper and rulers, you can of course easily prepare them for yourself.

For this chart field which measures 6 inches high (or for any other given size of chart) a general instructions-table can be used by which, without any figuring, you will know what

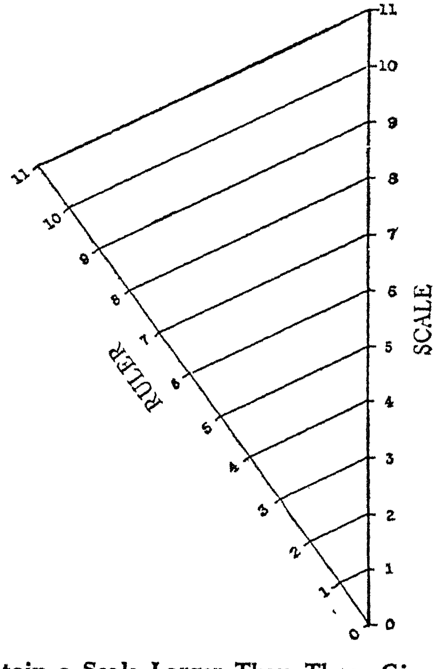


Fig. 168. To Obtain a Scale Larger Than Those Given by the Ruler.

type of ruled paper, what edge of the ruler, and what values or calibrations in the chart-scale you must use. In the accompanying table for 6-inch high chart-fields, it is assumed that you will position the peak of the curve about two-thirds of the way up the chart-field. You have therefore only to glance through your data and find out the amount of the peak or largest quantity in the series and with this figure in mind, consult the table and find the round figure therein which is nearest it, and proceed as for that round figure. The only

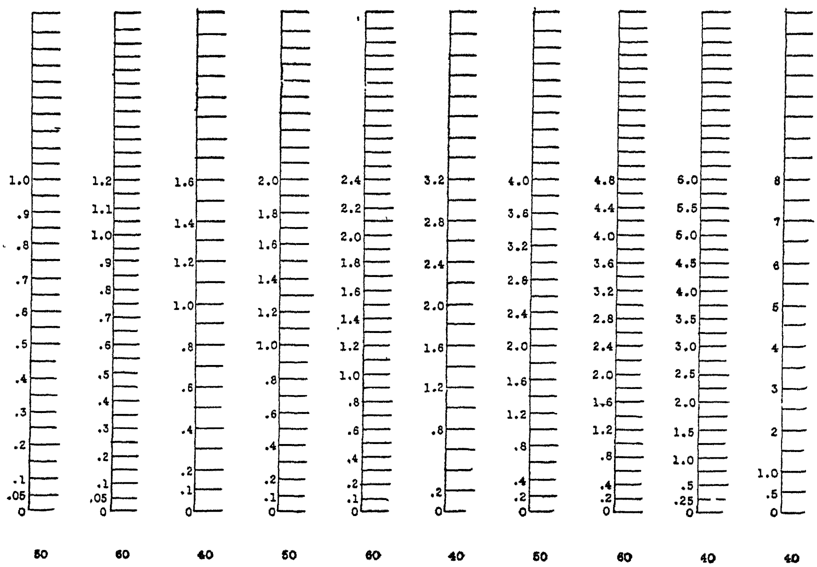


Fig. 169. Examples of Convenient Vertical Scales.
Reduced from Standard 6-inch Field.
Shows scales with 10, 12, 16, 20, 24, 32, 40, 48, 60, 80 and 100 at the distance of two-thirds of the height of the six-inch chart-field, by the use of the three rulings for 40, 50, and 60 sides of the ruler.

thing to remember is that the position of the decimal point does not matter. Your number may be .0003 or .3 or 3.0 or 30 or 3,000,000; you will always find its plotting instructions under the first "significant" digit, namely, in these cases, the figure 3.

These apparently arbitrary rules of thumb are justified only so long as they serve to produce the best results. Your real purpose is to show the data most clearly and simply, either to yourself or to someone else. The chart is a window,

ORDINARY PEAK IN ENTIRE SERIES	FIELD 6 INCHES HIGH					FIELD 8 INCHES HIGH			FIELD 10 INCHES HIGH		
	On Standard Fields										
	Type of Field to use	Scale-Figure first Abscissa	Value of Data to Plot	Rule to Use in Plotting	Plotting Value of Rule-unit	Value of Data to Plot	Rule to Use in Plot-ting	Plotting Value of Rule-unit	Value of Data to Plot	Rule to Use in Plot-ting	Plotting Value of Rule-unit
1.00 - 1.16	50	.05	2N	50	.1	N	20	.1	2N	30	.1
1.16 - 1.21	60	.05	N	30	.1	ditto			ditto		
1.21 1.26	ditto					ditto			N	20	.1
1.26 - 1.47	ditto					2N	50	.1	ditto		
1.47 - 1.54	40	.1	N	40	.1	ditto			ditto		
1.54 - 1.90	ditto					N	30	.1	2N	50	.1
1.90 - 1.95	50	.1	N	50	.1	ditto			N	30	.1
1.95 - 2.35	ditto					N	40	.1	ditto		
2.35 - 2.41	60	.1	N	60	.1	ditto			ditto		
2.41 - 2.52	ditto					ditto			N	40	.1
2.52 - 2.96	ditto					N	50	.1	ditto		
2.96 - 3.12	40	.2	N/2	40	.1	ditto			ditto		
3.12 - 3.80	ditto					N	60	.1	N	50	.1
3.80 - 4.62	50	.2	N	10	1.0	N/2	40	.1	N	60	.1
4.62 - 4.82	60	.2	N/2	60	.1	ditto			ditto		
4.82 - 5.00	ditto					ditto			N/2	40	.1
5.00 - 5.70	ditto					N	10	1.0	ditto		
5.70 - 6.10	60	.25	2N	30	1.0	ditto			ditto		
6.10 - 6.23	ditto					N/2	60	.1	ditto		
6.23 - 7.35	ditto					ditto			N	10	1.0
7.35 - 7.50	40	.5	N	20	1.0	ditto			ditto		
7.50 - 7.60	ditto					2N	30	1.0	ditto		
7.60 - 9.40	ditto					ditto			N/2	60	.1
9.40 - 9.60	50	.5	2N	50	1.0	ditto			2N	30	1.0
9.60 -10.00	ditto					N	20	1.0	ditto		
100 % Band	Increase 60-field by 2/3 Inch 60	2.5	2N	30	10.0 %	Increase Field by 1/3 In. N/2	60	1.0 %	N	10	10.0 %

TABLE OF SCALES FOR CHARTS

Fig. 170. Table for Vertical Scales with Engineer's Rules on 6, 8, and 10 Inch Fields.

as it were, through which the reader looks out upon an illuminating picture of the facts he is considering. Through this window he sees, if you like, a chain of mountains, whose height tells him the values or quantities he is considering. That he may see them to the best advantage, the window must be low enough for him to see the base of the mountain-range

and high enough for him to see at least some sky above the highest peak. In general, the best view of the mountains would show neither too much nor too little clear sky above. And if the window is crossed with a framework for small window-panes, he can further judge of heights by the criss-cross window-pane lines. Your curve is the silhouette of that mountain-range, your field the tiny window-pane outlines, and you, the chart-maker, must use your own judgment and artistic sense to place the reader's chair near or far, high or low, in front of that window, to give him the clearest view.

Chapter XVIII

PLOTTING-POINTS

A point can be defined or located by its co-ordinates. The co-ordinates of a point are the two mutually perpendicular lines which pass through it and at whose intersection the point is located. One of these lines is its abscissa, the other its ordinate. Neither of them need appear upon the paper; that is, they may both be imaginary, and it is therefore sometimes difficult to chart or plot a point precisely, or when plotted, to read its co-ordinates exactly.

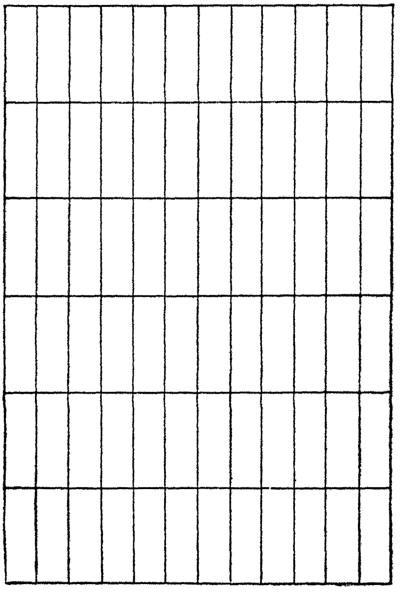


Fig. 171.

Here is a simple example. Suppose you have a chart-field on which each abscissa and ordinate represent a unit value, that is the abscissa and ordinate are numbered consecu-

tively one, two, three, four and so forth. Now suppose that you were to plot on the field the point represented by the co-ordinates, "x, 3½, y, 4½." You will look in vain for an intersection of lines with these values, because the co-ordinates of half units have not been drawn on the paper. Nevertheless, you can imagine the two co-ordinates, one of them half way between the ordinates of "x, 3" and "x, 4," the other half way between the abscissae of "y, 4" and y, 5." And at the intersection of these two imaginary co-ordinates you can plot the point.

This question of the precise plotting-point comes up very often in charts showing time by weeks, months, or years along the horizontal scale. Suppose you are charting the monthly steel prices in 1920. Down at the bottom of the chart you will place the time-scale with the words January, February, March, and so on under the ends of the vertical lines of the chart. Consider this scale carefully. What does it mean? It means that each unit of horizontal distance has been taken to represent one month and that all twelve horizontal units taken together represent a year, the year 1920. Now if a month were a single instant of time, it would be very simple. We would then plot the figures for each month or single instant of time on the particular vertical line which represented it. But as a matter of fact, a month is a long period of time with a great many different instants in it, all of which go to make up a single month, just as twelve months taken continuously make up a single year. In short, we are no longer dealing with single instants of time but with continuous periods of time. Yet on our chart the horizontal scale shows the number of single points representing these months. Something surely is wrong. Obviously we must find particular instants for points of time, to correspond with the points on the horizontal scale representing time.

There are two ways of doing this. The more scientific and accurate way is for us to seize upon the particular instant or point of time between months and represent these points of time by the points on our horizontal scale. The origin of the x-axis, or zero point on the horizontal scale will then stand for the beginning of the month of January. The first point on the horizontal scale will indicate the end of the month of January and the beginning of the month of February. The second point on the scale will represent the end of the month

of February and the beginning of the month of March, and so on. In this case we see that the months themselves are indi-

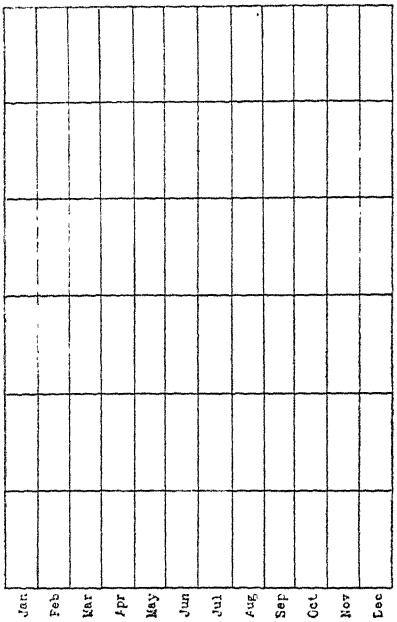


Fig. 172. To Plot Anywhere Between Ordinates.

cated on the scale, not by points, but by spaces between points. Thus if we wish to plot the figures for January as of the 15th of January, that is the middle of January, we will find a point midway between the zero and first upright lines, that is in the middle of the first space on the horizontal line. To plot a figure as of the end of the first week in January we will locate the point only a quarter of the way from the beginning of the horizontal scale to the first point on the scale, that is, a quarter of the way from the first of January to the end of January. And the scale itself, that is, the words "January," "February," "March," and so on, must be placed beneath the various spaces between lines, and not beneath the ends of the vertical lines themselves. This method enables us to distinguish prices at the various parts of each month, and is in general the more accurate method of scale calibration and point plotting.

The other method, however, is more convenient both for chart-maker and chart-reader. Let us assume that each month

has been condensed into a single instant of time and that the upright line or point on the horizontal scale represents only

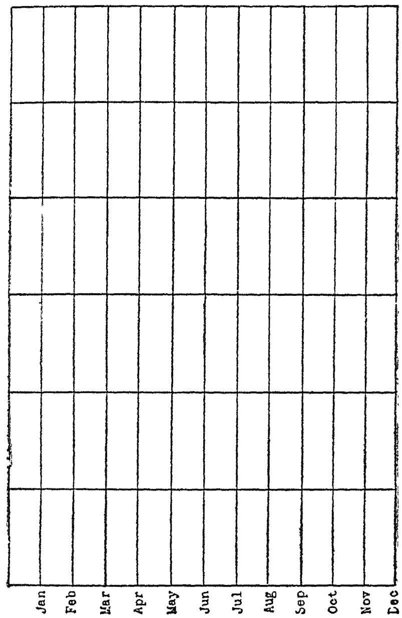


Fig. 173. To Plot Only upon Ordinates.

this single instant of time. What particular point of time in the month is in general the fairest one for us to choose to represent the whole month? Obviously the middle of the month or the middle of the fifteenth day. And when we plot the prices for January we assume that those prices are the average prices for the entire month and that it is fair to show them as the prices for this point in the month, namely the fifteenth of the month. In this case our scale figures will be written immediately underneath the ends of the vertical lines, that is, the word "January" will appear under the first ordinate and not under the first open space, the word "February" will appear under the second ordinate, and so on. Clearly this method is not scientifically so accurate as the first method. But as it is much more convenient, it is the ordinary method of plotting a time series. When you use it you must remember the assumption upon which it is based, namely, that the entire period has been condensed into a single moment or instant of time and shown as of that moment. Whether that single

moment be at the middle of the period or at some other time during the period will depend upon your data, and somewhere about the chart a memorandum should be placed showing what particular moment in the period shown, the figures and plotted points represent.

This second method can safely be used whenever the periods for which the data are charted, are uniform and equal periods of time. The method becomes very difficult and confusing

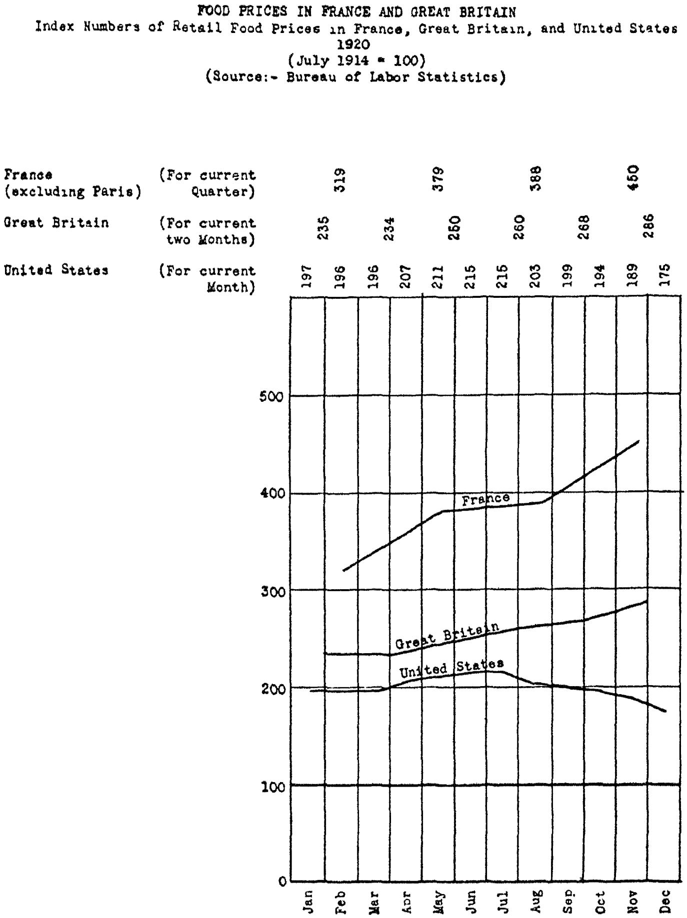


Fig. 174. Data with Different Intervals.

however, as soon as the time interval of the data changes. Suppose that a part of the year was represented by monthly

average figures and part of it by weekly average figures, and perhaps also a part by quarterly average figures. You would have to watch your step in plotting these various periods by the second method. But by the first method it is all smooth sailing, for it is easy to plot different points in the space when the spaces represent the months. It is also easy to plot a single point in the middle of three successive spaces (as for a quarterly period) by the first method. The first method, as has been said, is sound and logical and should be used whenever you are in doubt as to the plotting point on the scale. This means also that it should be used whenever the time intervals are not uniform and regular.

It is not necessary to say that one of the reasons why points should be correctly plotted is that the reader of the chart should be able to ascertain the values they represent directly from their co-ordinates on the chart. And the reader may have particular need for these values, not at the points plotted from the data, but at other points along the connecting lines which form the curve. The technical name for the process of locating points on a curve between given ordinates is "interpolation." Just as we can interpolate for points on the scale when plotting given points, so also we can reverse the process and interpolate for the data of points upon a given curve. In the foregoing we have shown how to interpolate for plotting points in making the charts. Let us now consider the reverse process of interpolating for data, either in the making or in the reading of a chart. It is to this process that the term interpolation is ordinarily applied.

Let us suppose that we have a chart compiled from data which is incomplete, that is to say, the months of June, July and August are missing. Our chart will show a curve extending over the first five months and the last four months of the year but there will be a gap during the three missing months. Yet we know that the phenomenon was in existence during these months. Therefore we cannot leave this gap vacant, for to do so would imply that the phenomenon had ceased to exist during that time. Now if we can make no guesses whatever as to the value which was missing, we should simply draw a dotted straight line across the gap connecting the two nearest known points of the curve—dotted or broken to indicate that data is missing and the curve for that distance is guess-work. If however, we can make a shrewd estimate as to the shape of

the curve across the gap, perhaps from a study of the same phenomenon in other years during the same months or from a study of similar phenomena during the same year, then we will not draw a straight line across the gap but will shape the cur ve over the gap in that way which we think most likely to be true. It would still be a good plan to use a dotted or broken line for these estimated or interpolated months. In any event, we have now assigned values to these months which were missing, by interpolation from a study of the surrounding ones which were known. The values of the interpolated points for these missing months can be read off from the chart and would form estimates for them with which we can even fill out the data record.

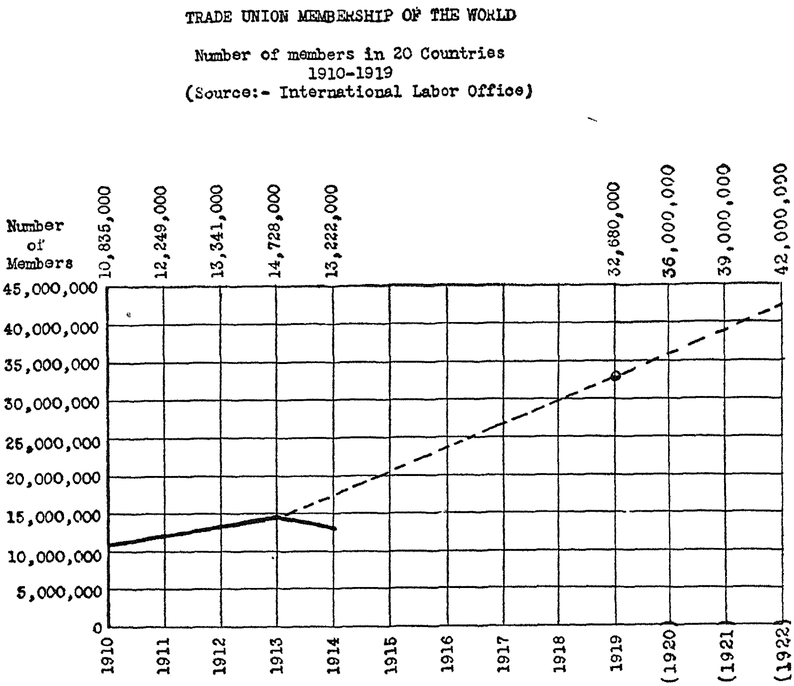
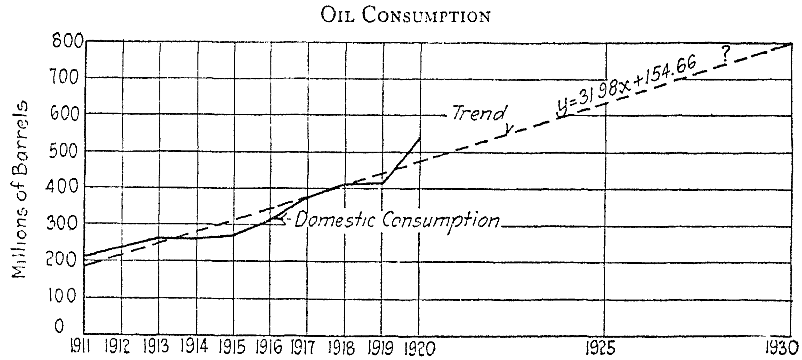


Fig. 175. Interpolation for the Period of the War and Extrapolation for the Years after 1919.

Interpolation for intermediate data from a completely known curve is also frequent. Thus if a curve shows the values at certain known points, we can easily secure the values at other in-between points by noting the points passed through

by the curve, that is, by interpolating for them. This guessing or estimating process can also be carried out beyond the limit of the curve to points lying outside of the range of the known data. A frequent example of this is the well-known process of



From Joseph E. Pogue's "Economics of Petroleum."

Fig. 176. Extrapolation.

Here it is the linear trend (having a mathematical formula) that is projected into the future.

extending a curve into the future to predict or forecast what will happen at a given time to come. The curve is simply projected to points outside of rather than inside of the range of its data. This latter process is often called extrapolation. The processes of interpolation and extrapolation are capable of very general application.

Chapter XIX

COMPOSITE CURVES

With the plain single-curve-chart, the reader of this book is now supposed to be thoroughly familiar. And as has been repeatedly indicated, the multiple curve-chart is formed by merely bringing together upon a single chart two or more curves, which may or may not cross each other. When two

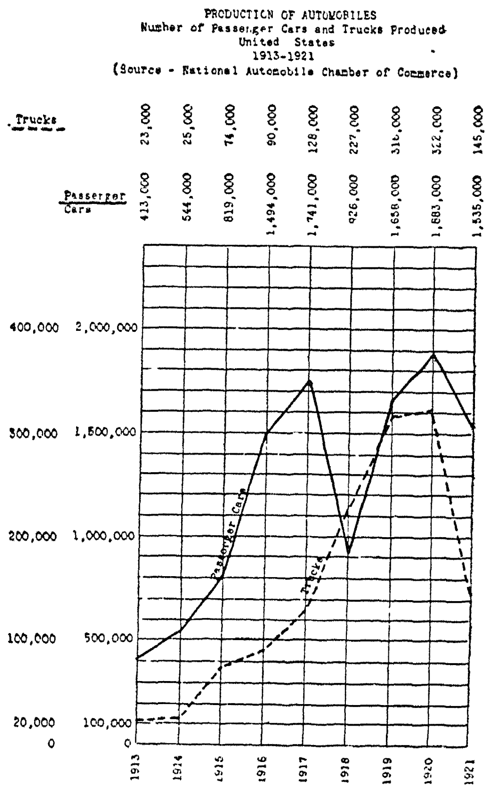


Fig. 177. Each Curve Has Its Own Vertical Scale.

curves cross at small angles, it is the better practice to distinguish them clearly, either by the use of different colors, or by the use of dotted or broken lines for one and full lines for the other. An older but slightly more confusing practice is to adorn one curve with small circles at its plotted points, another with small crosses, and distinguish other curves with double lines, wavy lines, lines with small cross-lines and sometimes lines of different thicknesses. In general, the best results are now secured by smooth lines, either colored or black, full or broken or dotted, but all of about equal thickness and visibility. Too many curves upon a chart are far worse than too

INVENTION AND WAR
Number of patents issued during the Civil War and during the World War
United States
1850-70 and 1903-20.
(Source: - U. S. Statistical Abstract)

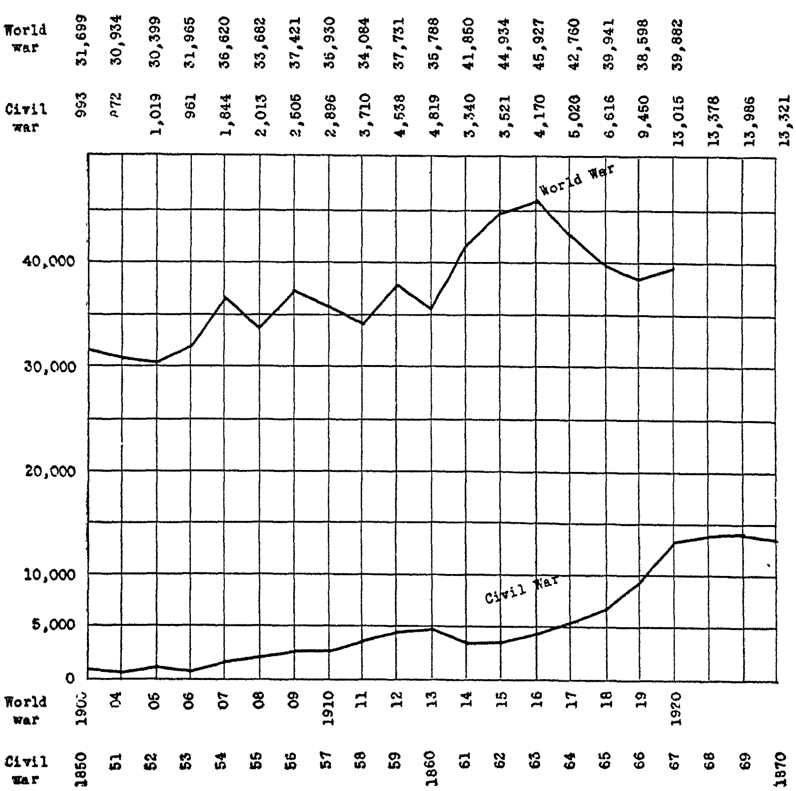


Fig. 178. Each Curve Has Its Own Horizontal Scale.

few, and except for special laboratory work a chart should not normally carry more than three or four curves. The value of attaching data is so great that it is unwise to dispense with data, and yet if too many curves are used, the data will bulk up disportionately. That a few curves can be easily distinguished without recourse to special adornments or thin and

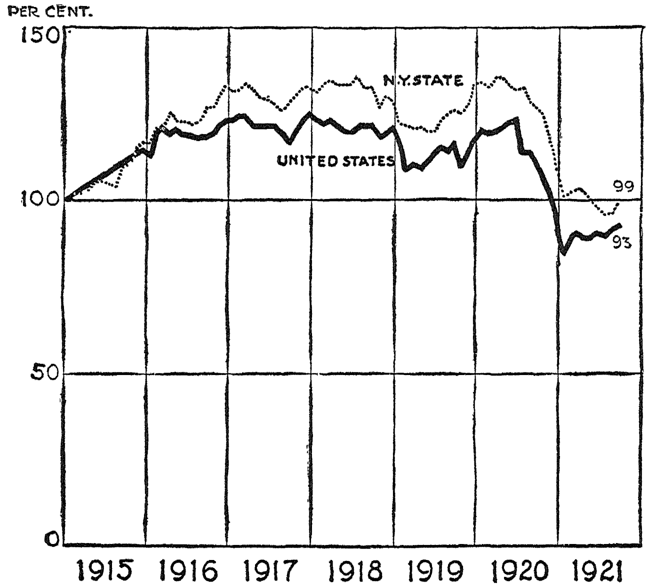


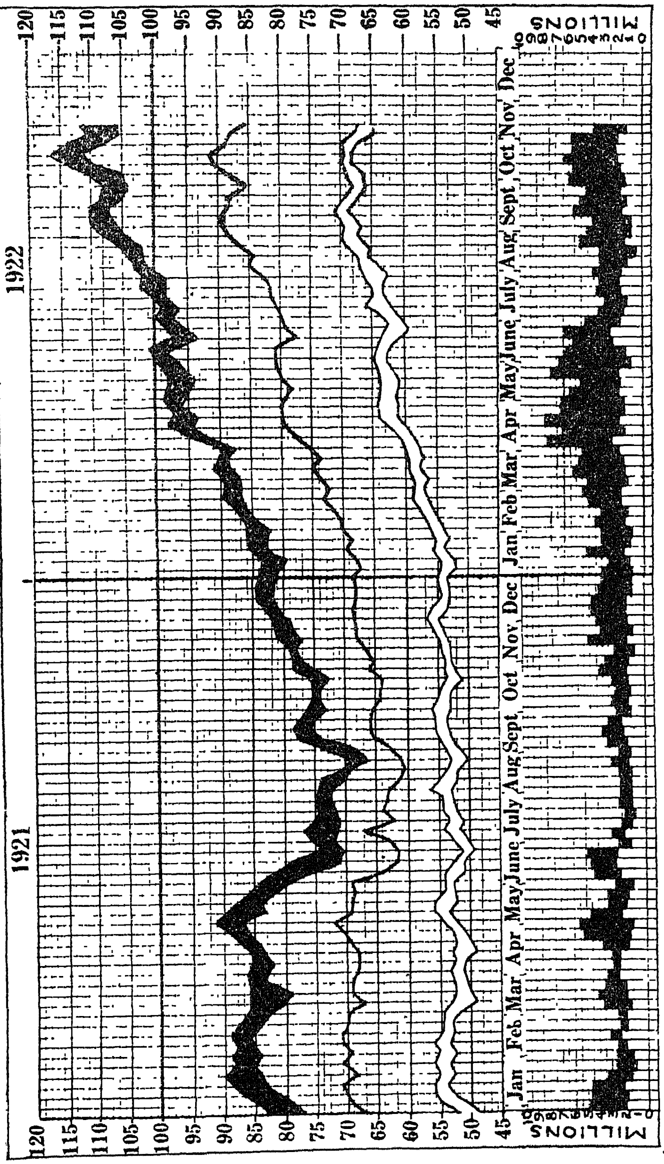
Fig. 179.

Number of persons employed in industrial establishments in New York State and in the United States (figures for December 1914 = 100%).—*Permission of Mr. Carl Snyder.*

thick lines, is obvious. The chief use for extra heavy or wide lines in curve-making should be for emphasis, as in the case of one curve for the average or total of the other curves.

A thorough knowledge of the curve-chart, however, requires at least a passing acquaintance with its sisters and its cousins and its aunts. We shall therefore hold a reception and introduce the most important of these. Beforehand, however, let us whisper a word in your ear about them. They are, none of them, such all around good fellows as the plain curve-chart. They are not so flexible and universal in their uses. Each answers excellently to certain limited types of data. We shall try to make you acquainted with the particular style of data for which each is best suited, as we meet them.

Consider the case of daily stock quotations on the exchange. For any particular commodity, a dozen different prices may be



In the upper portion the black line shows the closing average price of fifty stocks, half industrials and half railroads. The black area shows for each week the highest and lowest daily average price of the twenty-five industrials, and the white area the corresponding figures for twenty-five rails. In the lower portion the black area shows total weekly volume of sales and the white area weekly volume of the fifty stocks used in the preparation of this chart.

From The Annalist, N. Y.

Fig. 180. Zones Instead of Curves.

quoted in the same day. It is therefore customary to quote "highs" and "lows" as well as opening and closing quotations

in the stock market record. Now if we plot upon the ordinate for each day both the high and low quotations, and connect the low quotation points to make a curve for low quotations and similarly connect the high quotation points so as to make a curve for high quotations, we shall have two curves illustrating the same phenomenon, namely stock prices. The two curves show merely the extreme fluctuations of this phenomenon and the reader of the chart must understand that prices have ranged between these two curves. To make this situation obvious, let us shade the area between the two curves so as to make a zone. The shading can be done either with gray, with colors, with cross-hatched lines or with solid black or white, the co-ordinates being wiped out in the last case. This device conveys at once to the reader of the chart the idea that prices were not set at one figure alone, but varied considerably within the same day or period of time. The device can be used for any case of data covering maxima and minima

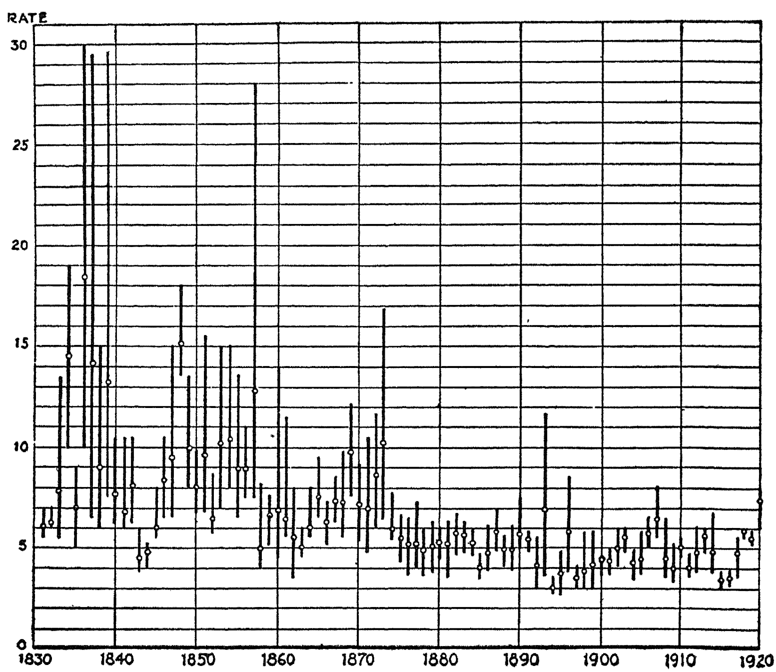


Fig. 181. An Excellent Form of Zone-Curve.

High, low, and average interest rates on commercial paper each year from 1831 to 1920.—*Permission of Mr. Carl Snyder.*

RETAIL FOOD PRICES
Index Numbers of the Bureau of Labor Statistics
United States
Jan. 1919-Dec. 1921
(Average 1913 = 100%)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
All articles	185	172	175	182	185	184	190	192	188	189	192	197
Sirloin steak	162	162	165	172	175	170	171	166	161	157	156	154
Round steak	175	174	177	182	187	181	185	177	170	168	162	161
Rib roast	165	165	169	175	178	171	169	164	158	155	153	153
Chuck roast	175	174	178	184	186	176	173	166	158	153	151	152
Plate beef	181	181	163	187	181	174	165	160	160	145	143	143
Pork chops	193	180	184	197	205	202	220	223	219	211	200	181
Bacon	217	206	205	212	210	212	215	214	206	196	189	186
Ham	199	195	191	197	203	205	211	212	205	198	188	186
Lard	211	203	211	223	246	254	266	266	242	228	231	221
Hens	188	186	193	202	204	200	197	196	194	189	184	184
Eggs	218	147	140	143	154	156	164	174	183	209	236	261
Butter	184	149	174	185	177	165	164	167	172	186	197	204
Cheese	201	186	183	190	191	192	195	197	195	192	195	196
Milk	175	174	166	169	167	169	169	174	176	180	184	188
Bread	175	178	175	175	176	177	173	180	180	180	182	182
Flour	200	203	206	218	227	227	227	224	221	221	224	231
Corn meal	207	200	197	200	207	210	217	220	223	220	220	220
Rice	159	164	154	154	154	159	168	178	190	199	202	203
Potatoes	188	182	171	182	194	224	282	294	253	224	229	253
Sugar	196	195	193	193	193	193	198	202	200	207	227	264
Coffee	117	123	126	129	136	143	156	160	164	159	164	164
Tea	127	126	129	128	128	129	130	130	130	131	131	127

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
All articles	201	200	200	211	216	219	219	207	203	198	193	178
Sirloin steak	159	160	161	170	171	182	192	186	185	177	171	166
Round steak	166	167	168	179	179	191	202	196	193	188	178	160
Rib roast	159	159	161	169	169	176	181	176	175	168	165	162
Chuck roast	158	157	157	166	166	174	179	172	170	162	158	145
Plate beef	152	152	150	157	155	157	168	154	152	147	146	138
Pork chops	178	180	186	206	202	194	208	219	238	238	210	167
Bacon	186	186	186	191	196	200	205	205	202	202	196	176
Ham	187	188	190	199	206	215	222	224	224	222	212	186
Lard	216	204	192	191	189	185	184	177	177	185	183	162
Hens	197	210	216	224	221	218	211	212	214	207	201	189
Eggs	240	199	161	163	153	156	166	184	206	234	250	268
Butter	194	190	196	199	187	175	177	175	179	180	181	162
Cheese	196	196	194	194	194	189	186	183	184	184	180	176
Milk	187	188	187	183	182	182	188	191	193	194	194	189
Bread	195	198	200	200	205	211	213	213	213	211	207	193
Flour	245	245	242	245	264	267	264	255	252	236	221	200
Corn meal	220	217	217	217	223	230	233	230	227	213	197	183
Rice	208	210	211	214	215	215	214	210	202	185	163	152
Potatoes	318	353	400	535	566	606	624	294	229	200	194	188
Sugar	324	342	340	367	462	485	482	416	373	263	235	191
Coffee	165	166	165	165	165	166	165	162	153	146	139	133
Tea	132	131	136	136	136	136	137	137	137	133	136	133

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
All articles	172	168	166	162	145	144	146	155	153	153	152	150
Sirloin steak	159	151	154	157	158	167	158	157	153	147	141	139
Round steak	163	153	157	160	160	160	161	160	164	148	139	138
Rib roast	157	148	152	154	153	151	146	147	144	139	136	136
Chuck roast	148	138	141	140	138	135	129	130	128	124	120	120
Plate beef	140	129	130	127	124	117	109	112	110	109	106	106
Pork chops	171	156	166	177	167	162	165	161	179	171	158	145
Bacon	171	166	165	164	161	159	160	162	159	153	147	145
Ham	180	179	181	183	181	182	190	197	191	180	170	165
Lard	141	131	124	116	106	103	106	115	113	109	105	101
Hens	200	201	203	202	194	181	182	185	179	175	168	168
Eggs	229	139	121	99	97	101	122	138	146	171	201	204
Butter	159	148	150	145	111	106	122	134	132	139	139	136
Cheese	176	171	176	169	143	133	133	148	148	149	151	149
Milk	183	173	171	167	162	160	157	161	158	160	161	158
Bread	193	189	188	184	177	175	175	175	171	170	166	163
Flour	203	197	194	179	173	179	176	171	170	164	155	152
Corn meal	173	167	160	153	150	150	147	150	147	143	140	137
Rice	137	121	113	106	101	101	100	101	103	107	108	107
Potatoes	176	153	147	135	129	159	200	247	235	206	188	182
Sugar	176	182	176	178	153	142	129	138	133	125	122	118
Coffee	129	126	125	123	121	120	120	119	119	119	119	119
Tea	133	131	131	129	129	126	127	127	127	127	127	124

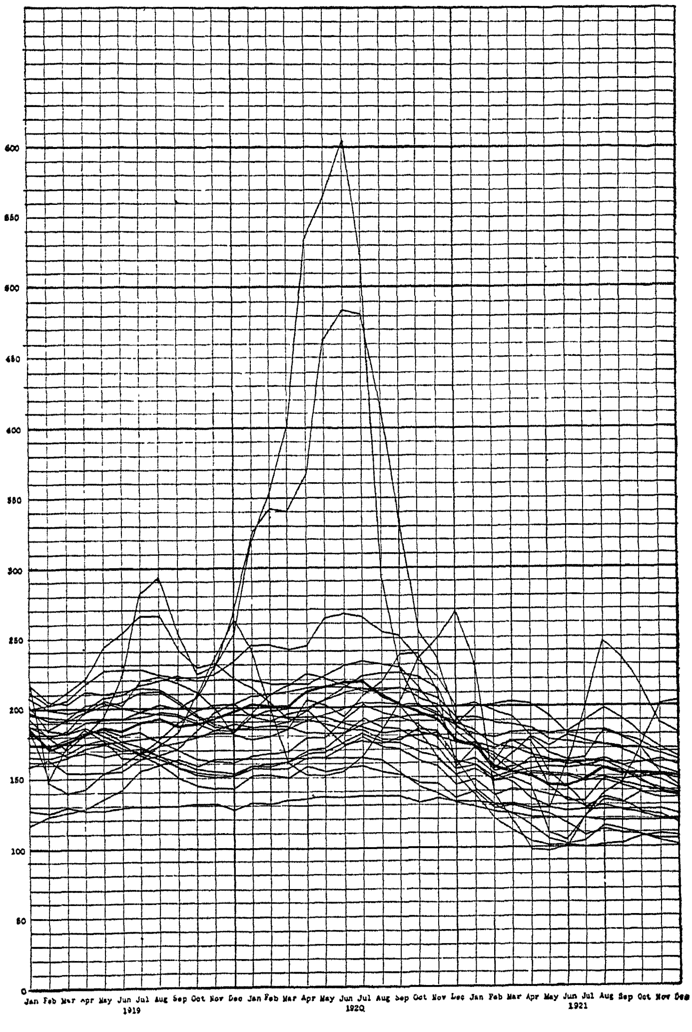
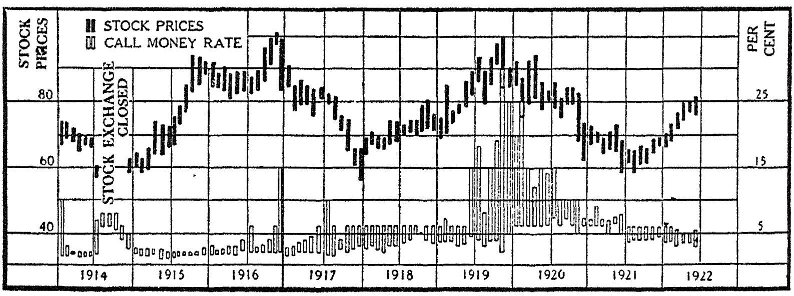


Fig. 182. It is Useless to Show All the Individual Curves.

for a single phenomenon. Climatic conditions, such as humidity at morning and night, or temperature at mid-day, midnight, and noon, or tidal variations, or business statistics such as the margins of profit from individual sales, and in general all data having a considerable range of variation for one and the same thing, at approximately one and the same time, can be shown by this method. This type of curve-chart is commonly called the zone-curve.

In a sense, the zone-curve is merely a short-cut for a large number of curves superimposed upon each other. In some cases you will have such distinct data that you could have prepared a large number of separate curves. If you put all these curves together upon a single chart, you will produce much the same visual result, so far as the reader is concerned, as you produce by means of the zone-curve, in which you merely plot maxima and minima and shade the space between them. Where the individual curves must be kept distinct



Permission of Standard Statistics Co.

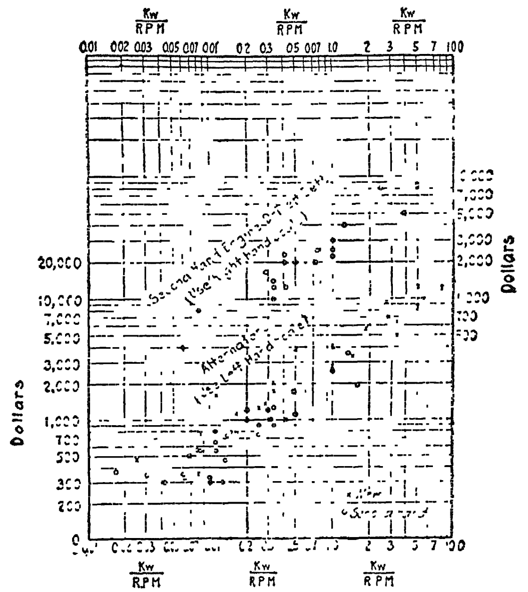
Fig. 183. An Excellent Adaptation of the Zone-Curve.

however, and compared with each other, of course the zone-curve is of no use. The zone-curve, therefore, is not a substitute for the multiple-curve chart.

We have said that the connecting lines between the plotted points which form a curve, imply a connection between the items of the data. We have said that this connection is established by the variable nature of the stubs, or x-axis scale. It is now time to let you into the secret that these plotted points do not need to be connected. The "gun-shot" chart is an example of a curve without curves. It consists entirely of plotted points. It is useful for cases of data secured by separate and often contradictory observations. Each observation

is plotted as a point but no connecting line can be drawn to other points and there is merely a large group of dots or plotted points extending across the chart and showing the result of various observations. This chart and its data differ from the zone-curve and its data in that there is no maximum or minimum known. Isolated points or dots on the chart may occur far outside of the general run or trend or zone of the main body of observations, such cases being due to freaks. errors, or other causes.

Gun-shot charts are essentially a research device. They are often intermediate steps between the first data gathered



From Leonard A. Doggett's "Cost per pound of Electrical Machinery," in the Electrical World, Oct. 2, 1915.

Fig. 184. A Gun-shot Chart.

and the final data reported. If we find after making a gun-shot chart of any particular observation, that all the points lie within a very narrow zone across the chart, we will be tempted to draw a line through this zone and so "fit a curve" to the plotted points. This is a sort of deductive reasoning by which we may often reduce a mass of data to a simple curve.

A popular form of the curve chart is the "staircase" curve, sometimes erroneously called a "histogram." The staircase curve is a direct throwback to the pipe-organ or vertical-bar

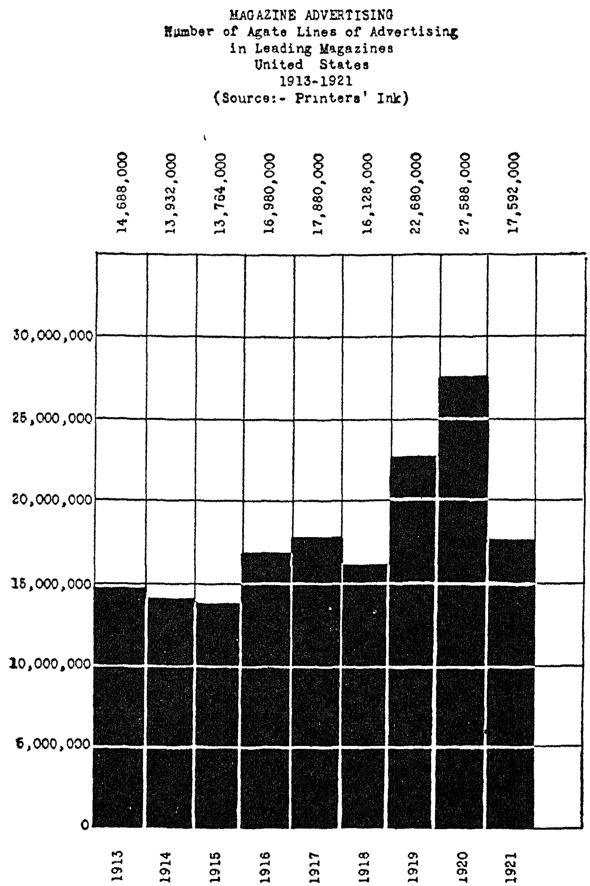


Fig. 185. The Staircase Curve is Almost a Bar-chart.

chart. If you will recall the original definition of a curve given in a previous chapter, you will remember that a curve may be defined as a line connecting the upper end of the bars in a vertical-bar chart. Now if you will make the bars wide enough so that they actually touch each other and will then draw the outline or silhouette of the upper ends, you will have a curve made of rectilinear lines always parallel to one or the other axes of the chart. This is the staircase curve. Whereas the ordinary curve represents the end of the bars by mere points and connects these points with straight lines, the staircase curve gives full value to the entire width of the bar. It is the precise silhouette formed by the bar-chart when the bars

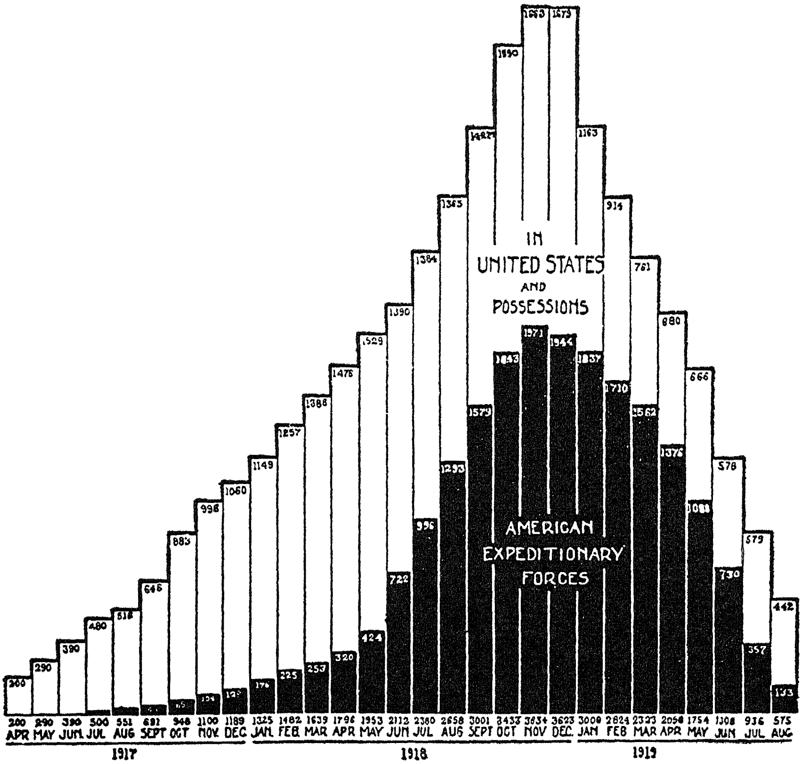


Fig. 186. An Absolute Compound Pipe-Organ Bar-chart or an Absolute Stair-cased Band-chart.

Thousands of soldiers in the American Expeditionary Force on the first of each month.—*Permission of Mr. Leonard Ayres.*

are packed close enough to come in contact with each other. As compared with it, an ordinary curve, directly connecting the midpoints of the ends of the bars, is called a smoothed curve.[1]

In some cases, the staircase chart is more accurate than the smoothed curve and its representation of areas lying between the base line and the curve, more accurate. A little study will show you that the connected-line curve has cut off little triangles from every bar whenever the curve descended

[1] Beside the staircased (or rectilinear) and the smoothed (or line-and-angle) curves, there is still a third which is of such doubtful value and great hazard as not to be mentioned here. It is the rounded curve, in which no straight-lines or angles occur, but all parts of the curve are rounded off by means of "French curves" or by free-hand drawing. It is discussed in the chapter on Frequency Curves.

and added little triangles to every bar when the curve ascended. These little triangles are sufficient to change the

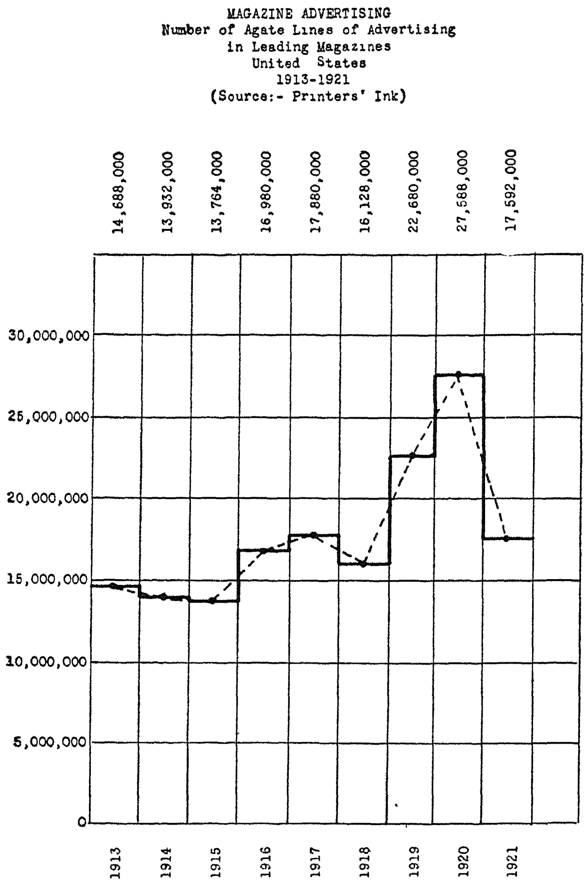


Fig. 187. The Smoothed and Staircase Curves Differ in Outline and Areas.

area lying between the curve and the base line, and bounded by the two ordinates about the plotted point, and when it is important that this area should be accurately shown, you can not use an ordinary curve but must use a staircase curve. At other times the staircase curve is less accurate than the smoothed one, for its abrupt changes of level give an impression of abrupt fluctuations in the phenomenon charted, which may be wholly unwarranted. The considerations governing the

comparative value of the smoothed and staircase form of curve are treated fully in a later chapter.[2]

The staircase curve is a popular form because it conveys at once to the average reader the impression of actual quantities between the base line and the curve. Readers who are confused by ordinary curves find less difficulty in understanding this chart. It is not, however, so useful as the ordinary curve because a number of these staircase curves cannot be satisfactorily put together upon a single chart. Their vertical portions will so often coincide that it is hard to distinguish them. The most that can be accomplished in the way of combining staircase curves is to put two or three of them together and use dotted, broken, and full lines to distinguish those which intersect.[3]

There is a certain type of data for which the plain ordinary curve closely imitates the staircase curve in its rectangular outline. This is the case of data in which the values remain ab-

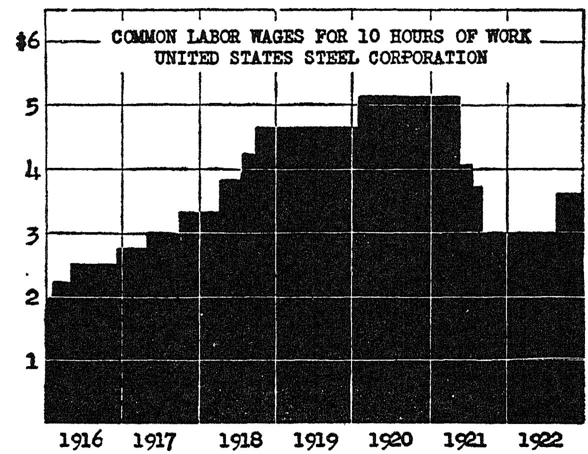


Fig. 188. Pseudo-Staircased Curve.

Note that so long as wages remain unchanged the curve must be a straight horizontal line and that when wages change the curve must be a straight vertical line. Hence the rectilinear form though truly a smoothed curve.—*Permission of Mr. Leonard Ayres.*

solutely fixed over a given period, and change only suddenly and abruptly. An example of this type of data would be the retail price of a single commodity, which after remaining at

[2] *Cf.* Chapter on Frequency Curves.

[3] *Cf.* Fig. 293, p. 333.

seventy-five cents for a long period of time suddenly and on a single day jumps up to one dollar, to remain there for another long period. Obviously to plot the 75¢-value by a dot at the beginning of that period and the $1.00-value by another dot at the time of change and connect the two by a direct line would give the impression of a gradual change extending over

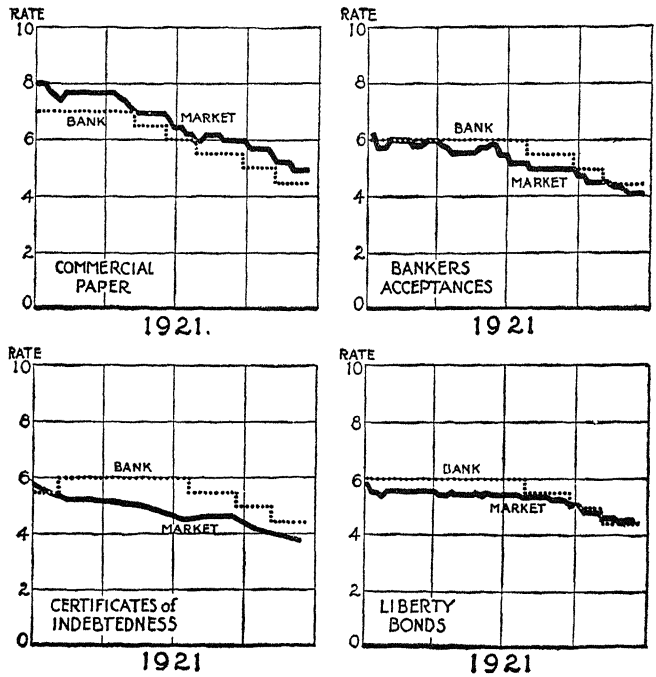


Fig. 189.

Open market interest rates at New York compared with the discount rates of the Federal Reserve Bank of New York. Open market rates shown are for prime 4 to 6 months commercial paper, prime 90-day banker's acceptances, certificates maturing in 4 to 6 months, and an average of the yields of 4 issues of Liberty Bonds and Victory Notes most frequently offered as security for advances. —*Permission of Mr. Carl Snyder.*

the entire period. It is therefore necessary that this curve should be perfectly level until the change takes place and then jump up to the higher level and remain there. The curve will then have a rectangular outline similar to that of the staircase curve, but the length of time or the length of the curve at any particular level is not regular and fixed. It is merely an accident that this picture has resulted in a curve with rectilinear outlines. It is not the same as the staircase chart.

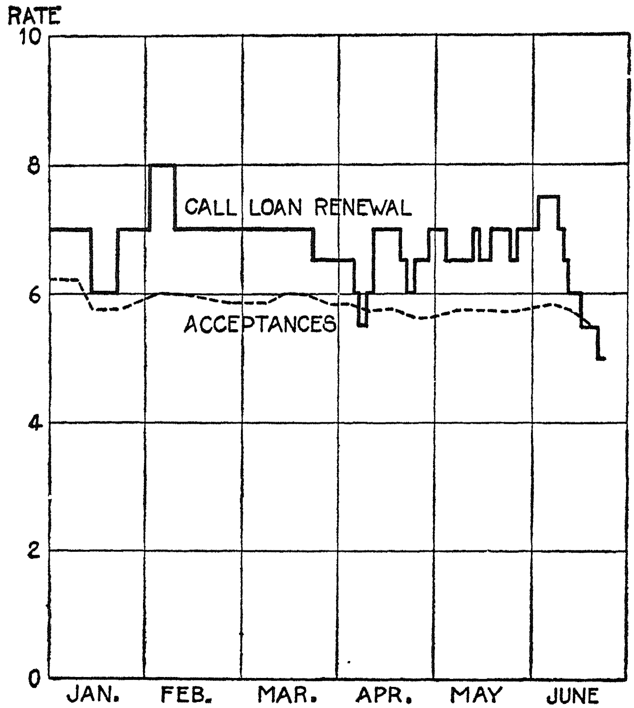
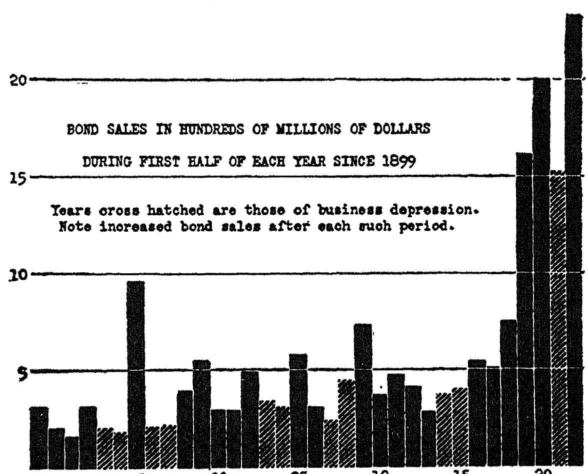


Fig. 190. A Pseudo-Staircased Curve.
Call loan renewal rate and prevailing rate on prime 90-day banker's acceptances at New York.—*Permission of Mr. Carl Snyder.*



Permission of Mr. Leonard Ayres.

Fig. 191. An Interesting Use of Shadings in a Band Chart, or Vertical Bar-Chart.

A gay and giddy member of the chart family is the "band-chart." Take up any of the ordinary curves which you have made and with a soft pencil shade the entire area under the curve. This vividly reminds the reader that the data is represented by the distance between the base line and the curve and not by the distance above the curve to the top of the chart, for it draws his attention forcibly to the lower part of the chart lying under the curve. You will remember the

THE FAMILY BUDGET
Divided as to Classes of Commodities
United States
1914-1921
(Figures as of December each year)
(Source:- Monthly Labor Review)

Total	103.0	105.1	118.3	142.4	174.4	199.3	200.4	174.3
Miscellaneous	21.9	22.9	24.1	29.9	35.1	40.4	44.3	44.1
Furniture and Furnishings	5.3	5.6	6.5	7.7	10.8	13.4	14.5	11.1
Fuel and Light	5.3	5.3	5.7	6.6	7.8	8.3	10.4	9.6
Housing	13.4	13.6	13.7	13.4	14.6	16.8	20.2	21.6
Clothing	16.8	17.4	19.9	24.8	34.1	44.6	43.9	30.6
Food	40.1	40.1	48.1	59.9	71.4	75.1	68.0	57.3

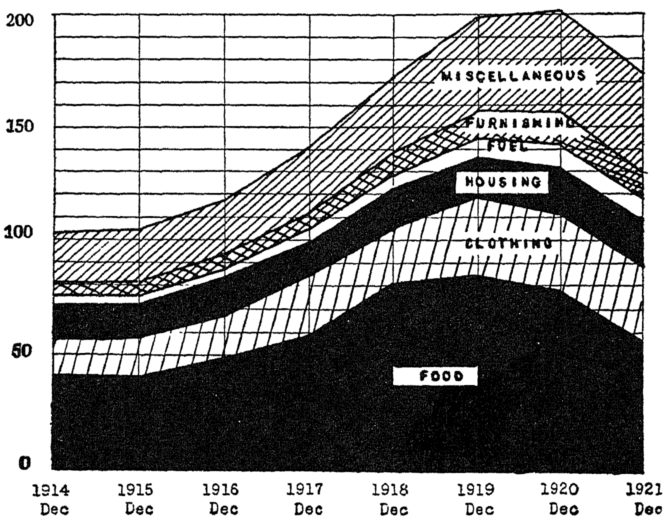


Fig. 192. The Curves are True Only for Cumulations of the Layers.

literal representation of quantities used in the bar-chart. And in fact the band-chart showing quantities by its shaded area

can be made in stepping form like the staircase chart as well as smoothed like the ordinary chart. The staircase band-chart is therefore even more of a throwback to the vertical-bar chart, or pipe-organ chart, than the staircase curve itself, because it has retained the shaded areas of the bars.

The band-chart becomes interesting when it is broken up into several bands running together across the page, each band representing a component part of the total amount under the curve. This is the band-chart proper, a series of layers or bands going across the chart which, when taken together, form a total whose fluctuations are shown by the curve of the top edge of the top band. This chart is sensational and interesting but of little precise value. You will find it hard, for example, to measure the width of any band except the lowermost. In fact the various curves which mark off these bands one from another have no value except that the lowest curve is a curve of one segment, the second curve is the curve for the total of the first two segments, the third is the total for the first three segments, and so on up to the top curve which is the total for all segments or the whole phenomenon.

FRENCH WOMEN-WORKERS DURING THE WAR

Proportion of Women to Total Employees in France
1914-1920
(Source:- Monthly Labor Review)

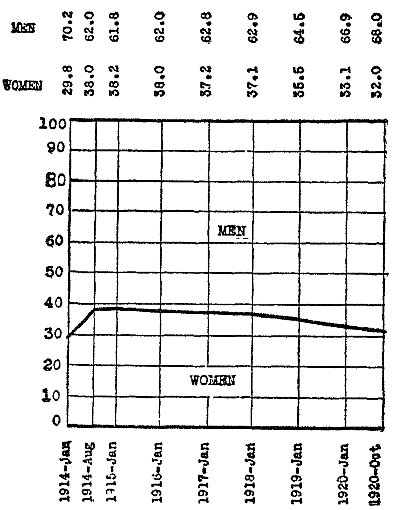


Fig. 193. A Relative (or Percentage) Band-chart.

And you will find that area conceptions are inevitably involved in this chart; the reader tries to measure the value of the various segments by the width of their bands. And unless staircased these areas will be extremely deceptive, the bands appearing to be narrower whenever the neighboring bands are moving rapidly up or down.

The most useful form of band-chart is the "100% band-chart." In this case the entire space between the zero or base-line and the 100% line is filled with various bands, each

CLASS ALIGNMENTS OF THE POPULATION
Divided into Capital, Labor, and Public
United States
1870-1910
(Source:- Arranged from Census by A. H. Hansen)

Capital	7.1	7.7	10.2	10.8	13.8
Public	58.2	53.7	48.1	45.4	41.9
Unclassified	8.1	8.2	9.3	8.5	6.0
Labor	26.6	30.4	32.4	35.3	38.2

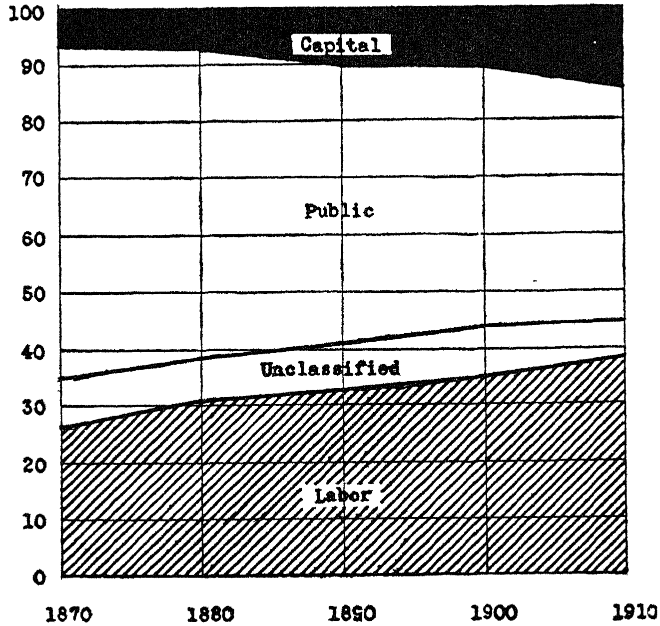


Fig. 194. The Smoothed Relative Band-chart.

THE NATURE OF EXPORT GOODS
Domestic merchandise exported classed as consumers' or producers' goods
United States
1910-1919
(Source:- U. S. Statistical Abstract)

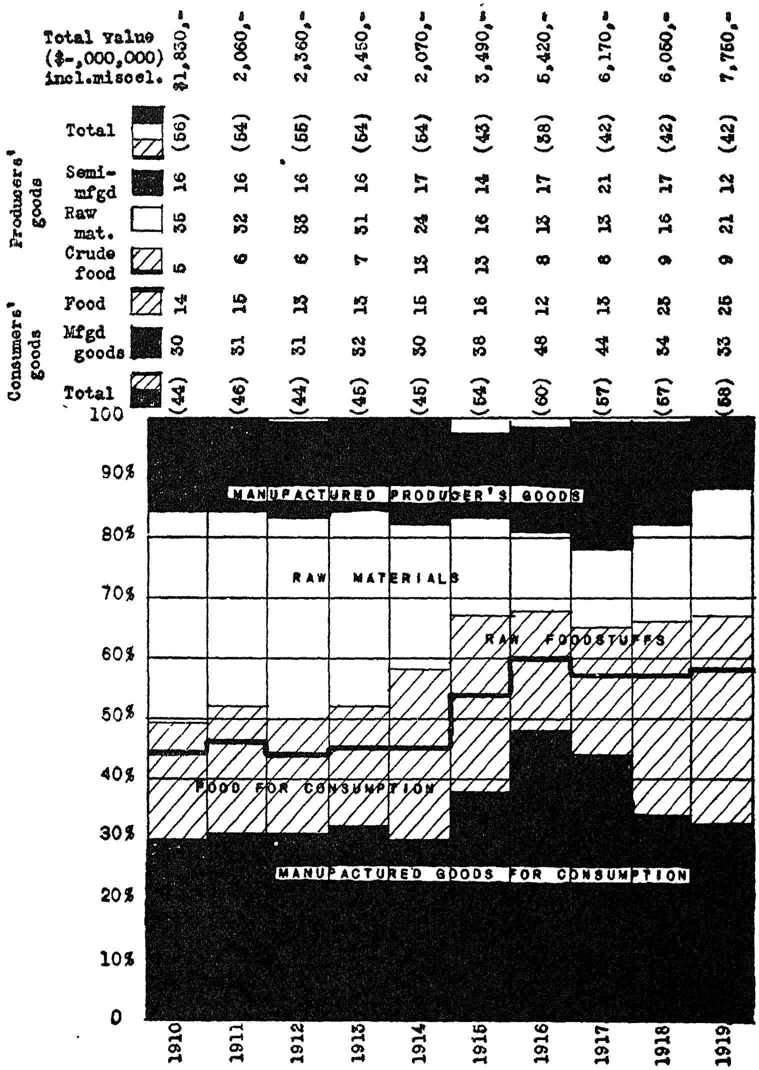


Fig. 195. The Staircased Relative Band-chart.

indicating a portion of the total or 100%. The fluctuation and changes of these bands show graphically the changes of

the component elements of this 100%. This type of chart is often used to show the changes in the distribution of cost and profit in an industry. The optional illusion of narrow bands when nearby bands are moving rapidly up or down is to some extent eliminated when the band-chart is made with stepping or staircase outline instead of smoothed polygon outlines. For

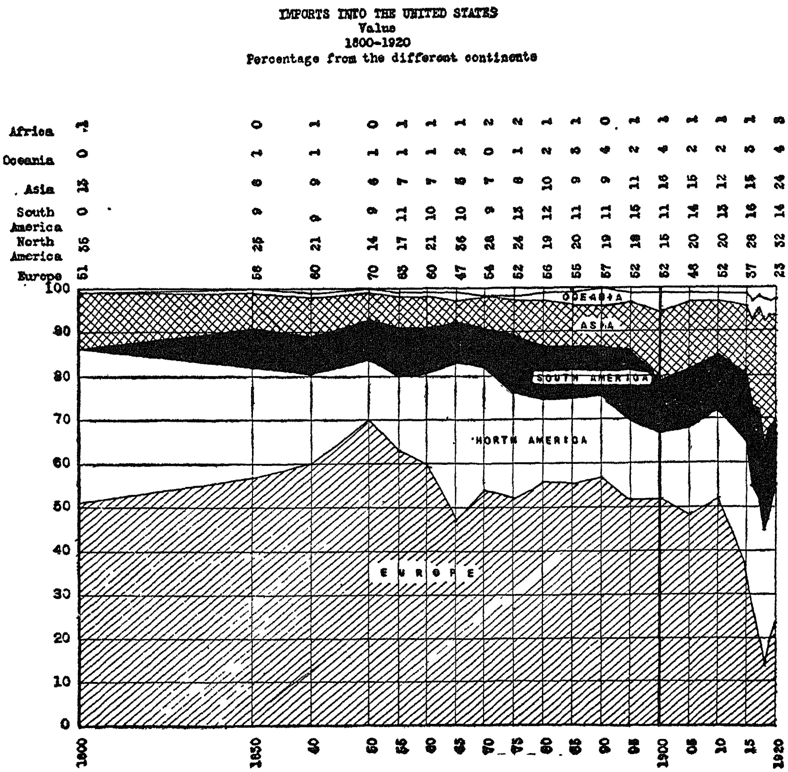


Fig. 196.

a simple presentation of the changes in the component parts of any phenomenon, this 100 per cent stepping band-chart is admirably suited. It is extremely popular in its appeal and does not suffer from the general disadvantage of staircase-curves because there is no question of superimposing other similar charts. It is the right way to represent cost components and other percentages to a general public, being within its narrow limits, safe, sound, and attractive. And the reader will notice that it is a form of curve which is well-nigh indis-

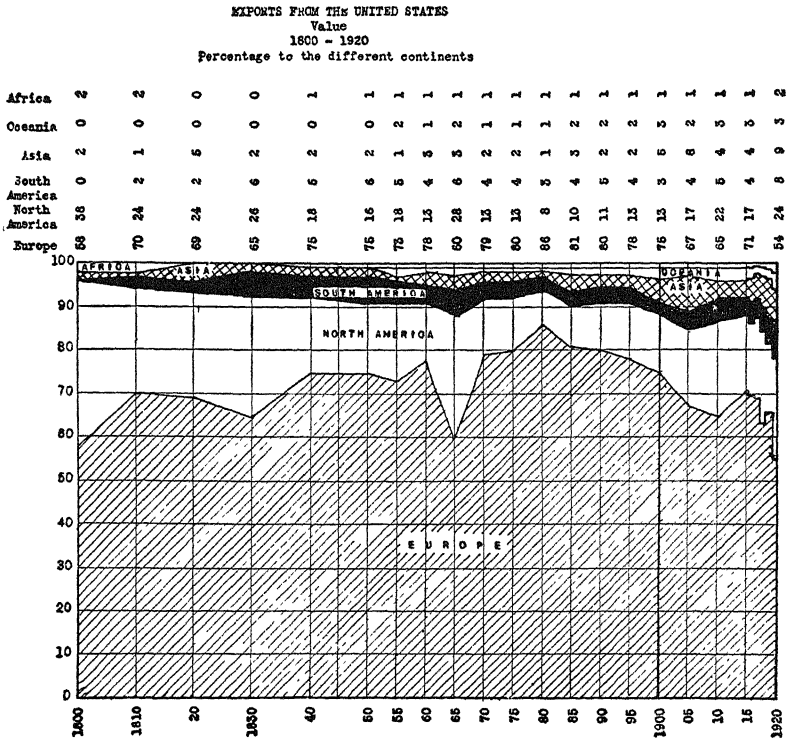


Fig. 197.

tinguishable from a bar-chart, being virtually a vertical-bar or pipe-organ compound, relative bar-chart, or in other words, a series of 100% bars set on end and brought into contact with each other.

In addition to the foregoing more or less distinct types of curves, there are also many and various possible embellishments which belong to the field of artistic rather than that of statistical endeavor. The object which is being charted may be pictured realistically and the picture shown at the end of the curve. Indeed, the same picture may be used frequently along the curve, or the picture may be modified to reflect the changes which are shown mathematically by the curve. Several different pictures may adorn as many different curves, and where one rises particularly high, it may be given a pair of wings or set in a balloon or aeroplane. By these and other fanciful ways, the imaginative chartmaker may make the appeal of his chart more vivid. But such measures are out-

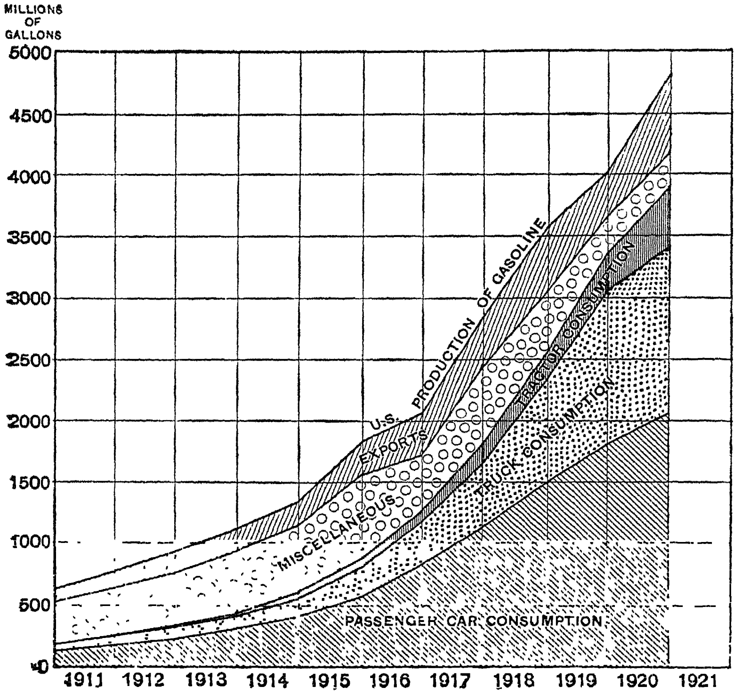


Fig. 198. An Excellent Band Chart (absolute).
Showing the consumption of gasoline by classes of uses.—*From Joseph E. Pogue, Economics of Petroleum.*

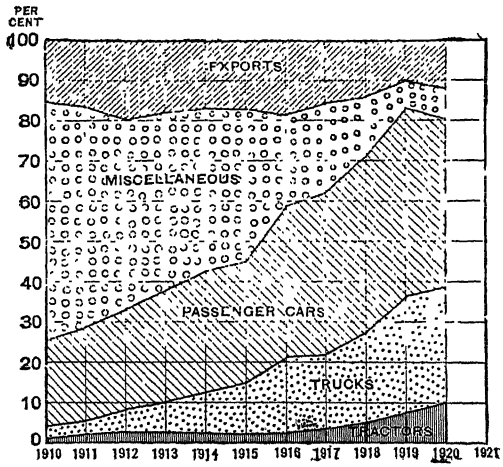
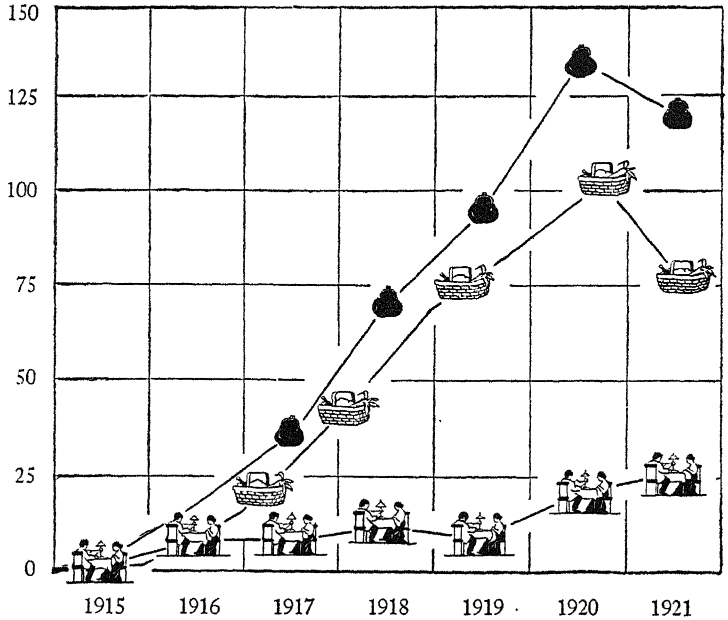


Fig. 199. The Relative Chart is Supplementary.
It is well to show data of this kind by two charts, the absolute and the relative or percentage distribution, and the latter can well be smaller.—*From Joseph E. Pogue, Economics of Petroleum.*



CHANGES IN THE STANDARD OF LIVING.
Index Numbers of Weekly Earnings in New York Factories, of the Cost of Living in the United States, and of the Living Standard ("Real Earnings").

Fig. 200. A Pictorial Curve.

side the proper scope of this book; the pictorial curve, like the pictorial bar-chart, is really intended for, and is useful for, popular consumption. We have come so far into the subject of mathematical charts that we shall hereafter have no time for purely pictorial effects. These may be left to the enterprise of the individual.

Chapter XX

HISTORICAL CURVES

Statisticians divide all series of figures into two groups. A series involving time, that is, a series for which different points or periods of time are the stubs or independent variable, they call a historical series. Other series in which time is not the independent variable, they call frequency series. This is a convenient classification for the chart-maker, and we can therefore divide all curves into historical curves and frequency curves. The historical curve has by some writers been called the "histogram," or "historigram," and the frequency curve

NEW INCORPORATIONS
Capital Invested in New Enterprises
Whose authorized capital equaled or exceeded $100,000
Principal States, U. S.
1919-1921
(Source:- N. Y. Journal of Commerce)
(Figures in millions of dollars)

	1919	1920	1921
Jan	492	2280	1243
Feb	324	1159	654
Mar	371	1376	955
Apr	516	1354	988
May	749	1418	601
Jun	1255	1323	676
Jul	1420	1260	282
Aug	823	941	580
Sep	1947	951	490
Oct	2364	1160	503
Nov	1341	896	368
Dec	1078	861	619

Fig. 201. A Historical Series.

the "pictogram," but these names have not been widely accepted in a precise sense.

In historical curves time is always the x-variable and must be plotted on the horizontal axis, its divisions forming the x-scale. In a previous chapter on plotting points the need for a precise scale has been discussed and the two methods of indicating periods of time, either by points on the scale or by spaces between points, have been dwelt upon. The reader is urged to review this section, as it meets a serious problem in the plotting of complicated historical data. The reader is also referred to a previous chapter on curve scales, in which the most useful forms and positions of the chart-field were explained.

The field of a historical curve-chart should be positioned very close to the righthand edge of the sheet of paper on

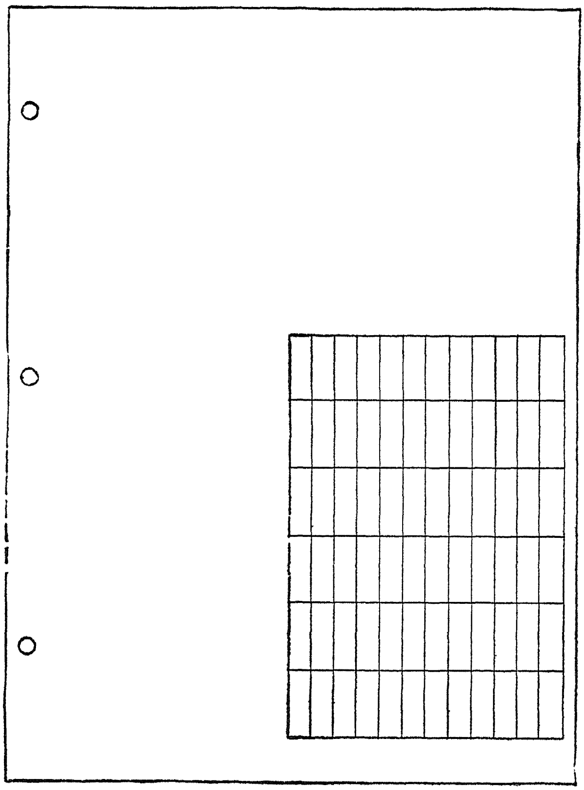


Fig. 202. Year by Months, Universal Ruling.

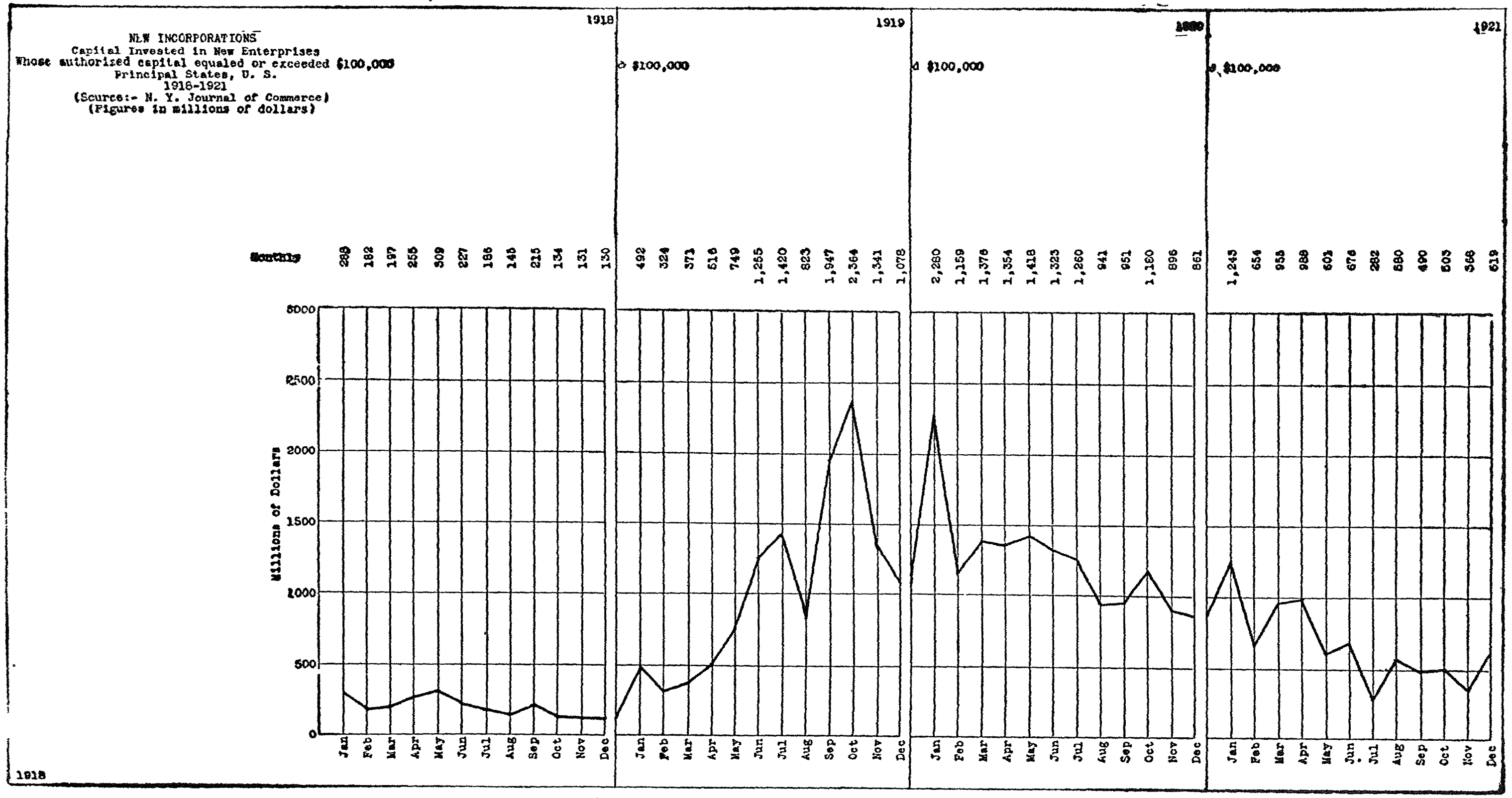


Fig. 203. The Individual Charts Combine Easily.

which it is drawn. There should be not more than a quarter or half inch margin between the chart-field and the edge of the paper on this side. The reason for this position will be clear to you the first time you prepare a series of historical curves in which the curve travels across sheet after sheet of paper through a succession of years or periods of time. By overlapping the charts (fanning them out) so that they are all visible, with only these narrow margins between them, the entire series can be made to appear as a single chart. In this form the entire series can be conveniently studied and economically photostatted. The narrow margin between each chart serves to break up the curve into its component periods without destroying its continuity. In this way a chart many feet long can be made upon ordinary sheets of paper without pasting them together and without inconvenience in filing or handling them. And from this long series, a single chart for a single period can be abstracted and individually compared with other individual charts. The narrow margin to the right of the chart may at first seem surprisingly inartistic, but it pays for itself in the flexibility of uses which it gives to the chart.

For a similar reason it is well to place the field of a historical curve-chart as low upon the page as possible. You must of course have room to enter the figures for the horizontal or x-scale legibly. This will rarely require more than three-quarters of an inch. By placing the field low upon the page, it is possible to compare curves for similar periods of time by laying one directly above the other, overlapping them vertically so that the two curves are both visible and their ordinates and time scales coincide. By this device, the reader can easily compare the seasonal or periodic fluctuations of two curves and at a glance detect the extent of their similarity.

In short, the field for a historical curve should be as close as possible to the lower righthand corner of the sheet of paper upon which it is drawn. This leaves a very large margin at the top of the page above the chart, in which should be entered the data of the curve. It also leaves a large margin to the left of the chart. This margin will be partly filled by the important vertical or y-axis scale or scales (if two or more scales occur on the same chart). But the chief use for the lefthand margin is that in it can be written the notes, comments or explanations which may be desired with the chart. If the

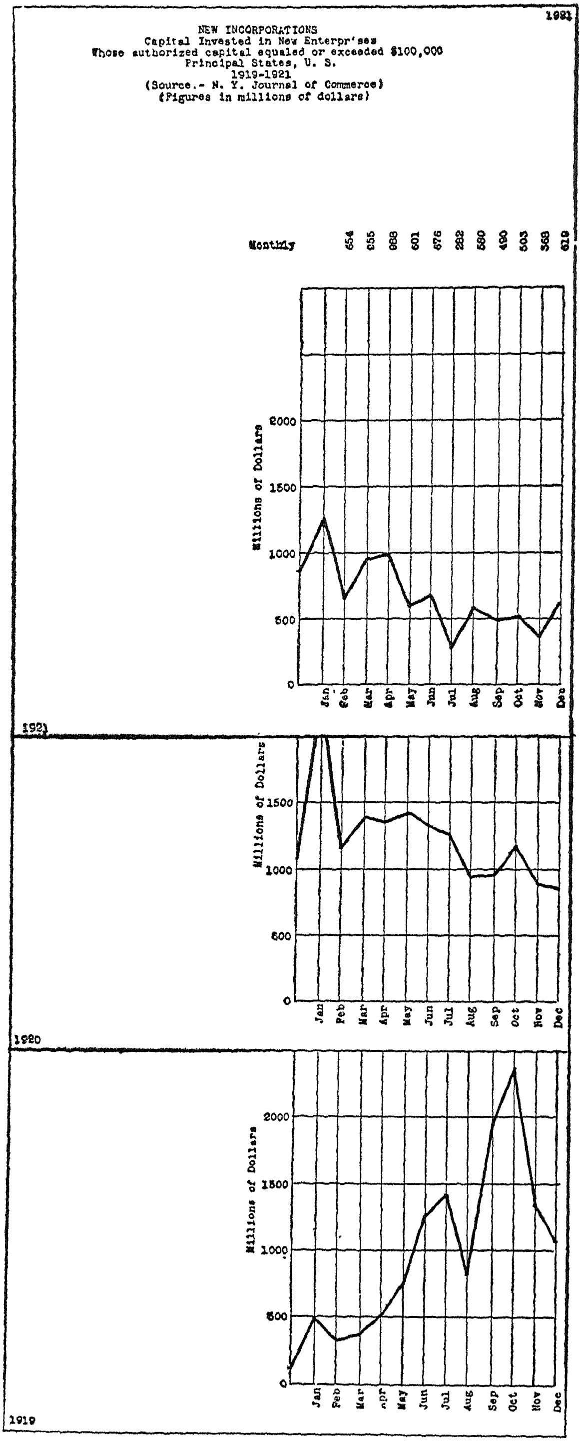


Fig. 204. Fanning Up and Down to Compare Seasonals.

sheets are to be bound in a loose-leaf holder or book, the binding edge will be on the extreme lefthand edge still further away from the chart. At the top of the page above chart and data, the title should be placed.

In historical curves, possibly more than in most, it is important that the data appear with the chart. It is important for the maker of the chart, for the curve is more easily plotted direct from the data, and the plotting checked for accuracy. It is important for the reader who is thereby enabled to either satisfy himself as to accuracy or to find any particular value without relying upon approximations more laboriously deciphered from the scale. The proper position for the data is, as has been said, above the chart, each value plotted appearing on line with the ordinate of its plotted point. Unless the data is extremely simple and brief and can, without crowding, be written horizontally, it is better to enter it vertically, writing or typewriting on edge, in the manner described in a previous chapter.

In historical curves, we have much use for a few simple mathematical and accounting phrases. The first of these is the "cumulative," or "total to date." When beside a column

NEW INCORPORATIONS
Capital Invested in New Enterprises
Whose authorized capital equaled or exceeded $100,000
Principal States, U. S.
1919-1921
(Source:- N. Y. Journal of Commerce)
(Figures in millions of dollars)

	1919		1920		1921	
	Monthly	Cumulative	Monthly	Cumulative	Monthly	Cumulative
Jan	492	492	2,280	2,280	1,243	1,243
Feb	324	816	1,159	3,439	654	1,897
Mar	371	1,187	1,376	4,815	955	2,852
Apr	516	1,703	1,354	6,169	988	3,840
May	749	2,452	1,418	7,587	601	4,441
Jun	1,255	3,707	1,323	8,910	676	5,117
Jul	1,420	5,127	1,260	10,170	282	5,399
Aug	823	5,950	941	11,111	580	5,979
Sep	1,947	7,897	951	12,062	490	6,469
Oct	2,364	10,261	1,180	13,242	503	6,972
Nov	1,341	11,602	896	14,138	368	7,340
Dec	1,078	12,680	861	14,999	619	7,959

Fig. 205. Simple Series and Annual Cumulations.

of figures showing the sales of your company, month by month, you place a second column of figures in which are entered the

NEW INCORPORATIONS
Capital Invested in New Enterprises
Whose authorized capital equaled or exceeded $100.000
Principal States, U. S.
1917
(Source:- N. Y. Journal of Commerce)
(Figures in millions of dollars)

Cumulative	312	663	1,084	1,523	2,008	2,431	2,924	3,386	3,643	4,024	4,396	4,617
Monthly	312	351	421	439	485	423	493	462	257	381	372	221

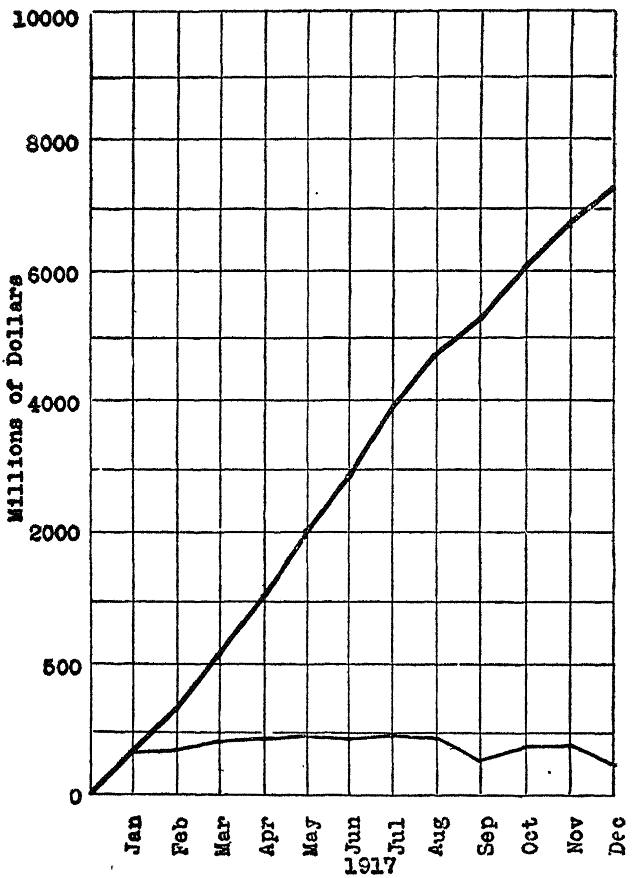


Fig. 206. Series and Cumulation Plotted With Same Scale.

total sales to date this year, your second column of figures is a "cumulative." The cumulative series begins with zero at the beginning of each period of time, that is, before the first value in the period, and is built up by adding each item to the previous cumulative until the final entry of the cumulative series is the total for the entire period.[1] The cumulative for the next period then begins with a new zero and similarly builds up to the total for the next period. And it is important to remember that at the end of each period, the cumulative has mounted to and equaled the total for the entire period. We deal here, of course, with periods of time (such as years) which contain a series of individual values for shorter periods, such as months.

Not all data can be cumulated. Economists make a distinction between "stocks" or "funds" and "streams" or "flows" of goods or money. A stock or fund of goods is something which can be considered in existence at a certain moment

1920	Jan	40.40
	Feb	42.90
	Mar	43.60
	Apr	43.68
	May	44.03
	Jun	44.80
	Jul	47.18
	Aug	49.12
	Sep	50.48
	Oct	49.21
	Nov	41.26
	Dec	36.96
1921	Jan	33.96
	Feb	31.48
	Mar	28.18
	Apr	26.98
	May	26.18
	Jun	24.72
	Jul	22.84
	Aug	21.98
	Sep	21.98
	Oct	21.98
	Nov	21.98
	Dec	21.98

WHOLESALE PRICES OF
BESSEMER PIG IRON
at Pittsburgh
(per long ton)
(Source - Bureau of Labor Statistics)

Fig. 207. This Data Cannot be Cumulated.

[1] By the forward cumulative described in the text we obtain "up to and including" figures. It is of course possible to cumulate historical figures backwards, obtaining "after and including" figures, but the step seems purposeless.—*Cf.* Secrist, Horace, *An Introduction to Statistical Methods*, pp. 232, 267.

or instant of time, while a stream or flow of goods is something which takes place during a given period of time. Figures of the latter, that is, stream or flow figures, can be cumulated. Obviously, if sales have continued throughout the year the sales for each month can be cumulated, that is, can be added together to give a total of sales for any period of several months or for the entire period of the year. On the other hand, the figures of a stock of fund or goods cannot be cumulated. In business, a common example of a stock or fund is the stock on hand or balance at any point of time. And obviously, if your balance was $3,000 on the first of January and $5,000 on the first of February, you cannot speak of your balance for the two months together as $8,000. It is not difficult to decide whether a series can be usefully cumulated. The use of cumulations or series of sub-totals is frequent in accounting.

The next mathematical conception is at present little used in ordinary accounting but is far more valuable for most analytical purposes than the cumulative. It is called the "moving total." To take the example given above, if beside your figures for the monthly sales of your company, you were to enter another column of figures showing the sales "for the last

NEW INCORPORATIONS
Capital Invested in New Enterprises
Whose authorized capital equaled or exceeded $100,000
Principal States, U. S.
1919-1921
(Source:- N. Y. Journal of Commerce)
(Figures in millions of dollars)

	1919	1920		1921	
	Monthly	Monthly	Moving Annual Total	Monthly	Moving Annual Total
Jan	492	2,280	14,468	1,243	13,962
Feb	324	1,159	15,303	654	13,457
Mar	371	1,376	16,308	955	13,036
Apr	516	1,354	17,146	988	12,670
May	749	1,418	17,815	601	11,853
Jun	1,255	1,323	17,883	676	11,206
Jul	1,420	1,260	17,723	282	10,228
Aug	823	941	17,841	580	9,867
Sep	1,947	951	16,845	490	9,406
Oct	2,364	1,180	15,661	503	8,729
Nov	1,341	896	15,216	368	8,201
Dec	1,078	861	14,999	619	7,959

Fig. 208. The Simple Series and Its Moving Annual Total.

twelve months," this new column would show the moving totals. Beside the January sales in 1921, you would enter the sales of the twelve months beginning with February, 1920, and ending with January, 1921. Beside the February, 1921, sales you would enter the total of sales for the twelve months beginning March 1st, 1920, and ending February 28th, 1921. It is easily seen that a moving total can be carried all through the year by simply taking the total sales for the previous year and successively subtracting the sales for the thirteenth month back, and adding the sales for the last month. Each figure in the moving-total series can be obtained by dropping off one month in the earlier year, and adding the corresponding month in the later year.[2]

The moving total is so useful that it has sometimes been enthusiastically described as the balance-wheel of commerce. When you are plotting the curves month by month, you are apt to find a considerable amount of monthly fluctuations, due in part only to normal seasonal conditions. In such cases, these perfectly normal seasonal fluctuations may hide

[2] The work-sheets for computing moving annual totals should always be designed to show similar months (or other parts of the cyclic period) together. This is done by arranging the periods in successive lines (with like months below each other) or in columns (with like months beside each other). Space should also be left for two other figures in each month: the first of these is the moving annual change, that is, the algebraic difference between the two like periods; the second is the moving total, that is, the cumulative of the moving annual change.

A second method is to arrange the monthly (or original) data in one long column. The comparison between like months may then be effected easily by means of a movable slip of paper with two slots or windows at the appropriate places to make the two desired months visible and hide all intervening months. The moving change and the moving total figures can then appear (each in a column) in two columns beside the column of original data. Totals should be taken directly from the original data at intervals for checking purposes.

If the cumulative also is being computed, still a different form of work-sheet is useful. In parallel columns the successive years should be tabulated, with each monthly figure on every fourth line down the page. The sheet should then be fed into a listing machine and the tabulated figures reprinted immediately below their entries. As they are listed, sub-totals should be taken on every record line below them, these forming the cumulative series. The sheet is then removed from the machine and the moving annual total entered by hand from a calculating machine (which adds and subtracts the proper months from the last totals and retains the results), these being entered in the remaining blank line for each month. This paper should be originally ruled with horizontal faints at listing machine intervals (of one-sixth of an inch) and with horizontal heavy lines every fourth line to separate the months, the page having 48 lines in all. As a further guide, vertical faints can be ruled in at listing machine intervals (one-sixth of an inch). If not specially printed up to order, the cheap cross-ruled paper with lines every sixth of an inch can be used. These details all tend to make checking up for errors very simple, and largely eliminate mistakes.

or obscure the true trend of the business, or at least make the determination of the trend more difficult. But you can easily tell whether the general trend of your business is upward or downward by plotting the curve for the moving annual total. The moving total series appears to flatten out the seasonal fluctuations and respond only to the true movements of the trend. In fact the moving total is sometimes called, even by statisticians, the "trend."

The moving total need not be annual, but can be computed for any given period of time. Thus we may have a moving 24-months total, or a moving 5-year total. In any case we have again periods within periods, as in the cumulative. The most usual form is the moving annual or 12-months total, for ordinarily in business there is a certain amount of normal monthly or seasonal fluctuation which repeats itself every year. These annual seasonal fluctuations are naturally swallowed up in a total for twelve months, for such a total always includes every month in the year.[3] The moving total is in general an excellent device for smoothing out the wrinkles and wiggles in a curve and reducing the curve to a simple regular trend-line. It should be used whenever the cycles of fluctuations appear to be of regular and uniform length or periodicity.

A word of caution is necessary about the plotting of a moving total. Strictly speaking, each item in the curve should be plotted in the centre of the period which it covers. Thus, the plotting point of each figure in a moving annual total series would normally be midway between the ordinates of the sixth and seventh months covered by the figure. The entire period of the total being one year, each point should be placed in the middle of the year which it represents. In this case, the moving

[3] Whenever the period of the annual cycle is not regular, as in crops and temperature cycles (one period of 12½ or 13 months, the next of 11½ or 11 months) it is well to follow Professor Secrist's suggestion of a thirteen-month moving total. This has the further advantage of centering the moving total figure precisely upon a monthly one (the seventh) instead of midway between two monthly ones (the sixth and seventh).

The same advantages are much better secured by an average of an eleven-month and a thirteen-month moving total, both centered on the same months. This may be called a "taper-smoothed" eleven-thirteen-month moving total, as it gives full weighting to the central eleven months and half-weighting to the terminal months (first and thirteenth). It will be seen that this precisely corresponds (in the average) with the periodicity of eleven to thirteen months. The taper-smoothed eleven-thirteen-months moving total is easily computed from the twelve-months moving total, as it is the two-months moving average thereof. Of course, a longer taper can be used if desired. The test is smoothness of the resulting curve.

total curve will begin five and one half months after the beginning of the curve of individual months, and will end five

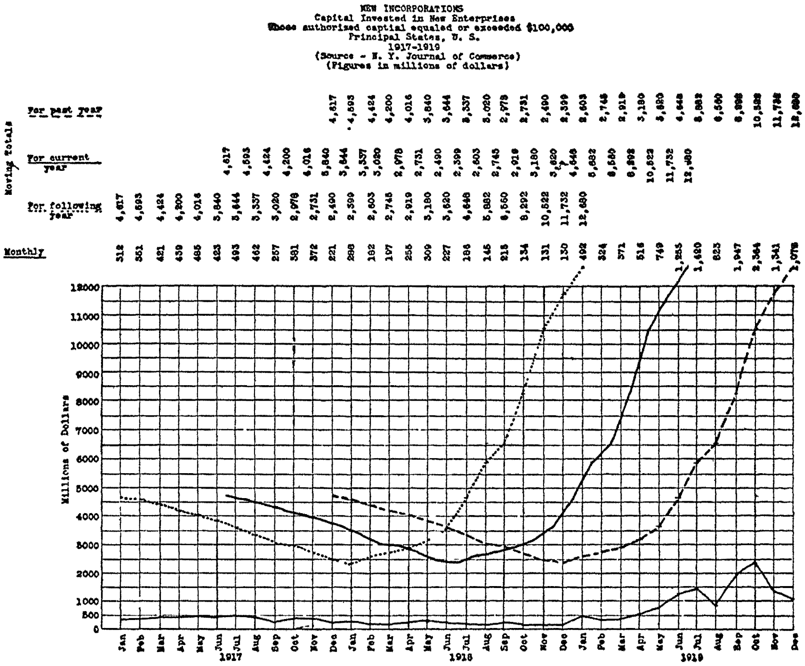


Fig. 209. Three Positions for the Same Moving Total.

and one half months before the end of the monthly curve. The moving annual total may then be called a "total for the current twelve months." For certain purposes, however, you may desire to have the moving total end on the same ordinate as the monthly curve. This can be done by plotting each item at the end of the year which it represents. It has the advantage of giving a more up-to-date appearance to the chart, but it is now necessary to label the series moving "total for past twelve months." On rare occasions, you may desire to place the moving total at the beginning of its period, in which case it becomes a "total for the following twelve months." When comparing trends, however, between various items, it is important that these moving totals be plotted at similar points

NEW INCORPORATIONS
Capital Invested in New Enterprises
Whose authorized capital equaled or exceeded $100,000
Principal States, U. S.
1918
(Source:- N. Y. Journal of Commerce)
(Figures in millions of dollars)

Moving Total for Previous Year	4,593	4,424	4,200	4,016	3,840	3,644	3,337	3,020	2,978	2,731	2,490	2,399
Moving Total for Current Year	3,644	3,337	3,020	2,978	2,731	2,490	2,399	2,603	2,745	2,919	3,180	3,620
Moving Total for Following Year	2,603	2,745	2,919	3,180	3,620	4,648	5,882	6,560	8,292	10,522	11,732	12,680
Monthly	288	182	197	255	309	227	186	145	215	134	131	130

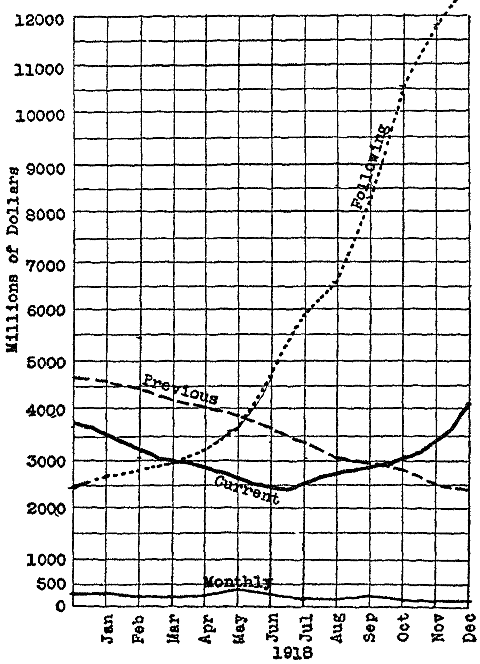


Fig. 210. A Detail of the Last Figure.

in their periods, else an unwarranted lag will appear between the fluctuations of the two charts.[4]

Similar to the moving total is the "moving average." The moving average is merely an average secured by dividing the

	1920		1921	
	Moving Total	Moving Average	Moving Total	Moving Average
Jan	14,468	1,203	13,962	1,163
Feb	15,303	1,276	13,457	1,120
Mar	16,308	1,361	13,036	1,086
Apr	17,146	1,429	12,670	1,054
May	17,815	1,485	11,853	987
Jun	17,883	1,490	11,206	934
Jul	17,723	1,478	10,228	853
Aug	17,841	1,488	9,867	822
Sep	16,845	1,403	9,406	784
Oct	15,661	1,304	8,729	727
Nov	15,216	1,268	8,201	684
Dec	14,999	1,250	7,959	663

NEW INCORPORATIONS
Capital Invested in New Enterprises
Whose authorized capital equaled or exceeded $100,000
Principal States, U. S.
1920-1921
(Source:- N. Y. Journal of Commerce)
(Figures in millions of dollars)

Fig. 211. Moving Annual Total and Average Series.

moving total by the number of items which compose it, that is, a moving monthly average is secured by dividing a moving annual total by twelve.[5] The moving average has the great advantage of lying at about the same height on the chart as the curve of individual periods (e.g. months), from which it is

[4] Needless to say, the only accurate picture of events (as regards time) displayed by the moving-total or average curve is that shown by the curve of the current twelve months or other period, that is, the curve of points plotted at the centers of their periods. (*Cf.* Chapter on Plotting Points, *supra.*) This is the true smoothed curve. At all other positions, the curve has been arbitrarily "lagged" forward or backward.

[5] The work-sheet for the moving average is the same as that for the moving total already described, save that an additional space must be left in each month (in the second, or columnar, method, it would be an additional column) for the moving average, which is derived from the moving total by dividing the latter by twelve (annually, or by whatever the number of items be which go to make up the total) (Obviously in the taper-smoothed eleven-thirteen-month moving total, the two terminal months have only half weight and the total is still of twelve months.)

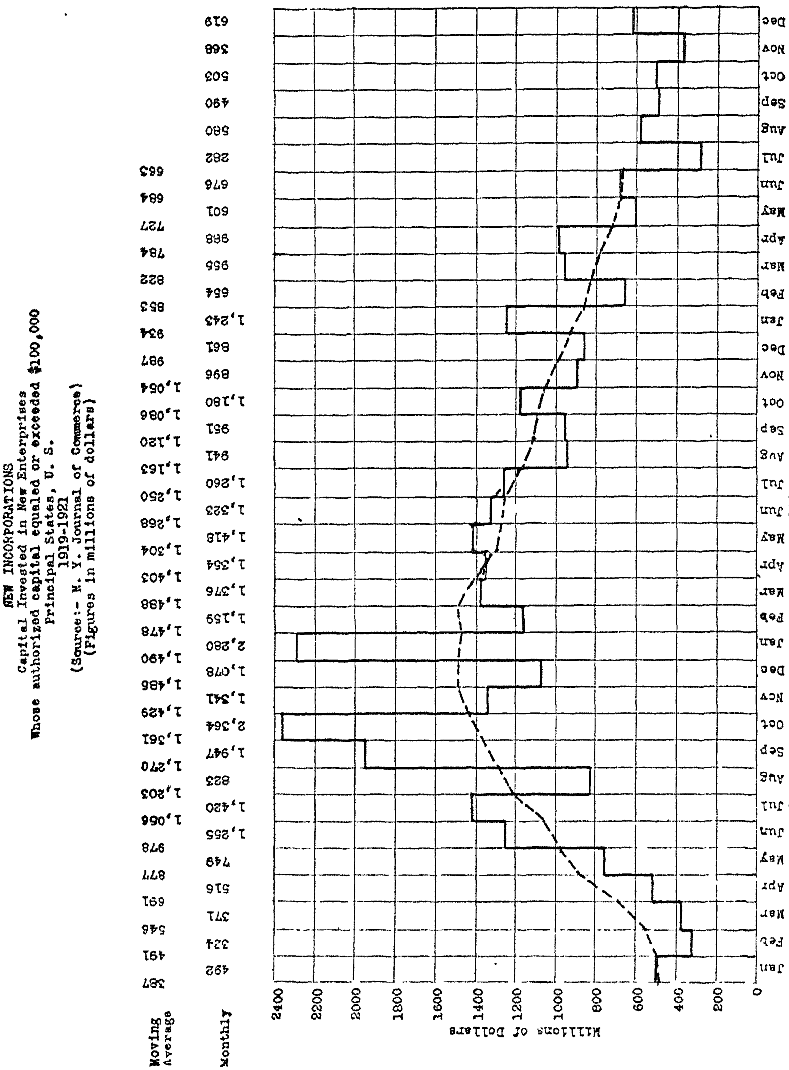


Fig. 212. The Moving Annual Average Gives the Trend.

derived.[6] And for this reason, you will find it an even better method of smoothing curves and showing their true trends than the moving total.[7]

---

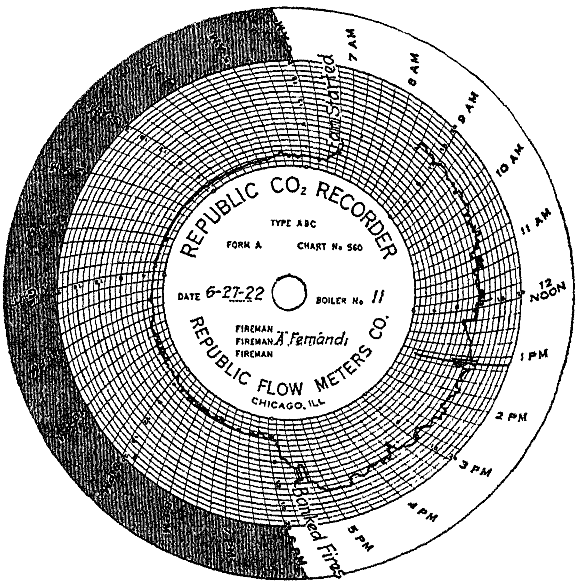
[6] The moving average has another great advantage over the moving total, in that it can be given variable period-lengths to conform to cycles of variable lengths. This is not possible with the moving totals.

[7] The moving totals and averages are also sometimes called "progressive" totals and averages.

Chapter XXI

CYCLES

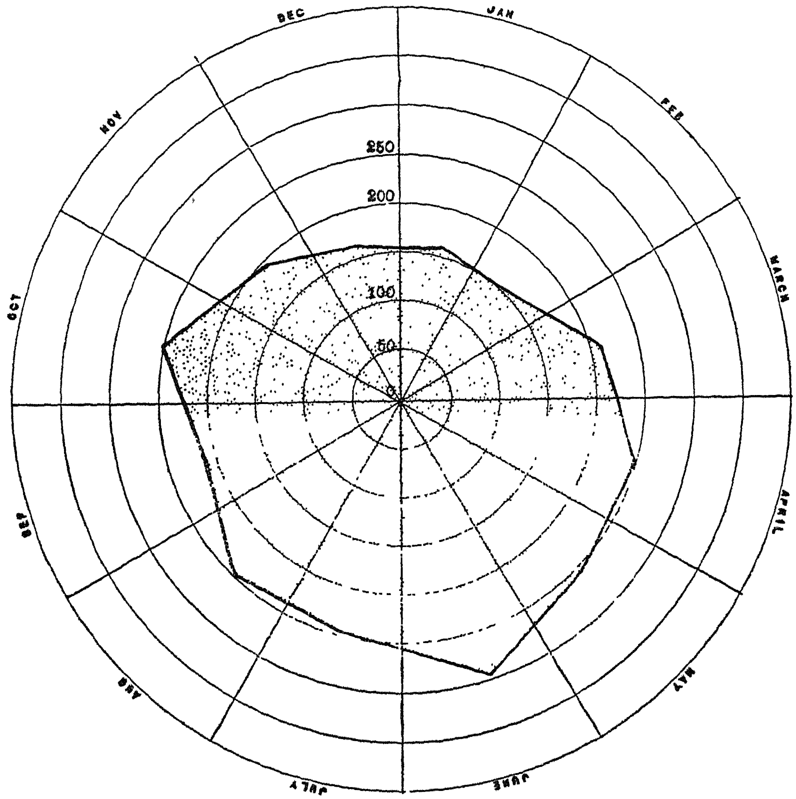
Long ago, or was it yesterday, there were neither automobiles nor aeroplanes, and the streets were frequented by a cheery and wholesome class of persons, who conveyed themselves about on two-wheeled contrivances called bicycles. In deference to our age, the reader will permit us to pause and sigh a moment over this happy retrospect. Sometimes in the circus, a contemporaneous antiquity, trick riders rode one-wheeled affairs with perilous skill. Needless to say, the rims of these wheels were as smooth and regular as the circumference of the average clock-chart.



Permission of Mr. Walter N. Polakov

Fig. 213. The Mechanical Cyclograph.

Now a clock-chart—if you have forgotten your early chapters in this book—is round like the face of a watch. Radiating lines or radii take the place of ordinates, and concentric circles or rings take the place of abscissae. Hence the chart can be used for the display of recurrent data—that is, historical data which after a certain period of time repeats itself. Your only care must be that the period of the cycle be adjusted evenly and wholly in one complete revolution about the circle. Then the curve of the data will meet at the



SEASONAL FLUCTUATION IN BUILDING OPERATIONS
AVERAGE YEAR, 1910-1920, IN 25 STATES
(Source:— F. W. Dodge Co.)

Fig. 214. As a Chart, This is Worthless.

two ends of the period, forming a continuous and endless curve.[1]

Had the trick cyclist in the circus used a wheel the rim of which followed the uneven outlines of this curve, he would indeed have had a bumpy ride. And a line drawn on the wall behind him, following the shadow of his head, would mark the same curve plotted on a chart-field of plain co-ordinates. Study the curve as so plotted in the ordinary way, and you will see that once every so often the wiggles or fluctuations

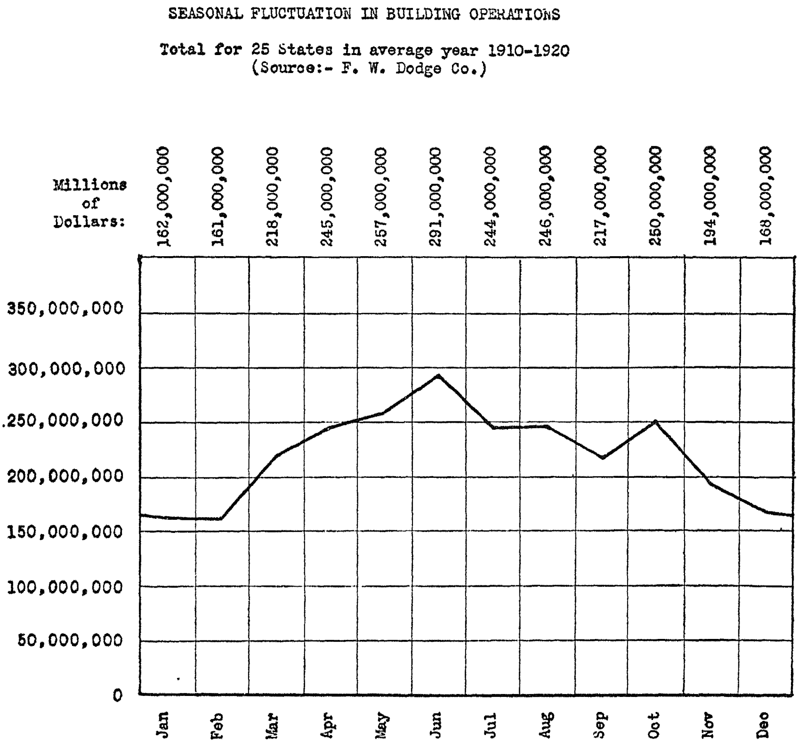
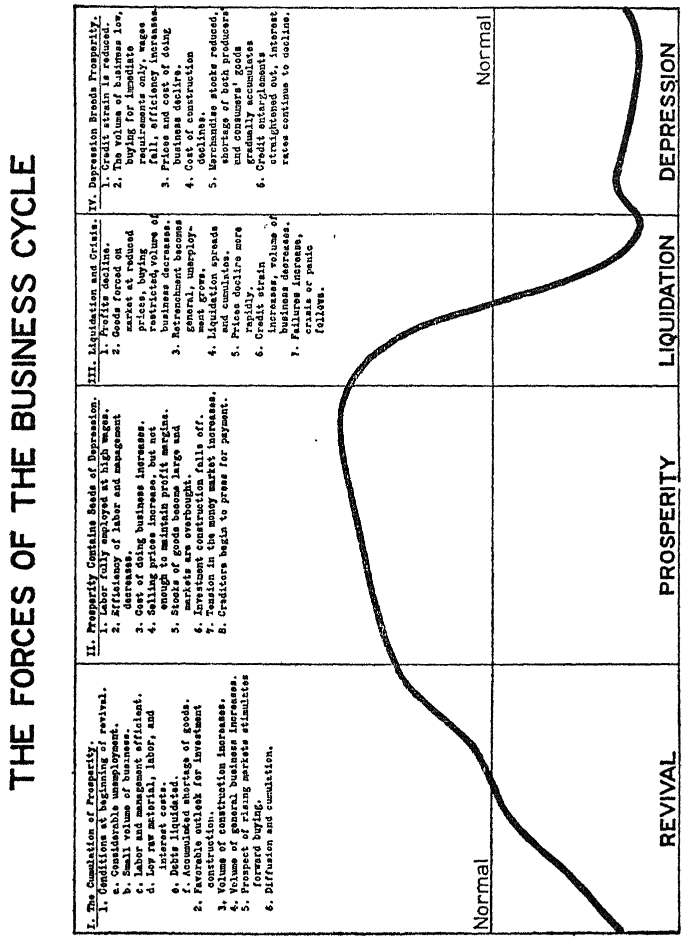


Fig. 215. The Rectilinear Co-ordinates Are Much Better.

[1] The clock-chart is similar to carp, the fish which is properly prepared by throwing it away after it has been cooked. When the clock-chart has been well and carefully drawn, it is ready for the waste-basket. For this reason, no detailed discussion of it or its polar co-ordinate field is entered into. The only case in which the clock-chart is a justifiable product is the case of automatic mechanical charts or cyclographs. These are parts of recording machines for temperature, pressure and the like, in which a fountain-pen at the end of a pointer leaves an inked record or curve upon the rotating disc underneath it. They are graphic records, but not otherwise useful charts.

of the curve repeat themselves, like the digits in a recurrent decimal fraction. So whenever in a historical curve you detect the same or similar fluctuations repeating themselves throughout the curve, you are justified in suspecting that in the repeated unit or part you have a cycle. It is an important function in statistical work to detect the presence of these



Permission of Mr. Malcolm C. Rorty.

Fig. 216. A Free-hand and Imaginary Picture of the Business Cycle.

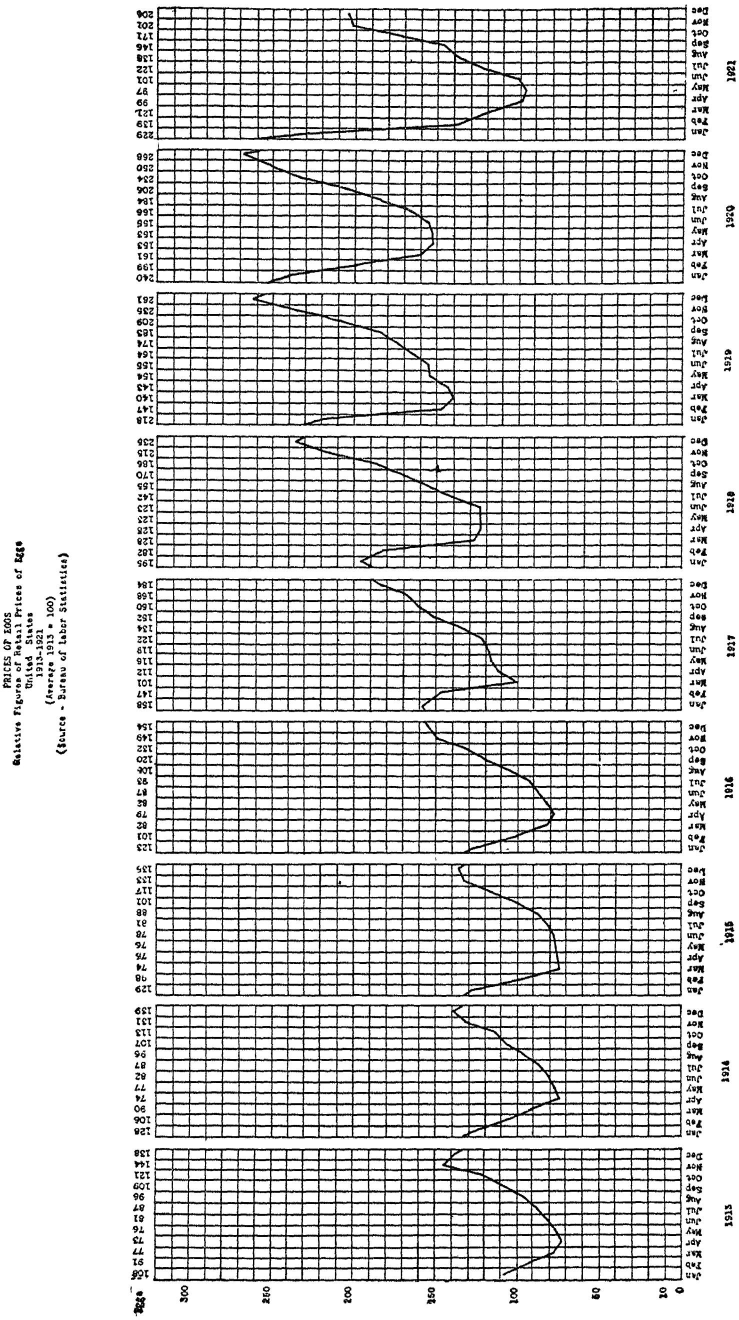


Fig. 217. Cycles of Slightly Varying Lengths.

cyclic fluctuations, and to be able at will to remove them. The subject has already been touched upon in the last chapter.

The time between the commencement of one cycle and that of the next, is called the period of the cycle. This period may be short or long, according to the nature of the data, a very frequent short cycle being that of 24 hours or one day. In business statistics, there are cycles longer than a year. Some investigators have found evidence of business cycles in which eras of general prosperity, depression and crises repeated themselves every four, eight, or even twenty years.[2] In meteorological and astronomical sciences, cycles of dry and wet weather have been found to last thirty-three years and of warm and cold weather about a hundred years. The most important cycle in most business statistics is the annual one, of four seasons or twelve months. The student of business statistics can almost always assume that he will find more or less of an annual cycle of seasonal fluctuations. Sales may repeatedly rise in spring and fall and decline in summer and winter. Production may then fluctuate somewhat earlier in the year, anticipating the changing demand. If production is uniform, ware-housing cycles will appear in order to absorb the surplusage in low selling periods.

It is not often that the recurrent cycles are identical either in the shape or the height of their curves. Such variations may be due of course to incidental and insignificant causes, but in general studies of broad trade or economic movements, they are often given a more fundamental importance, as being significant of real changes in the phenomena studied. The problem then is to isolate these variations in the cyclic fluctuations, that is, to eliminate from the series its seasonal cycles and retain its significant changes. To the series which remains after the removal of cyclic fluctuations, the name of "secular fluctuations" if often given. And we may therefore look upon the original series as being a combination of two different sets of forces, or movements, which we call, respectively, cyclic change and secular change. Either or both of these elements may be the object of your study and it is important that you should be able to determine them easily.

---

[2] The literature on this subject is considerable. In particular, the student should refer to:

Mitchell, Wesley C., *The Business Cycle*, and

Moore, H. L., *Economic Cycles: Their Law and Cause.*

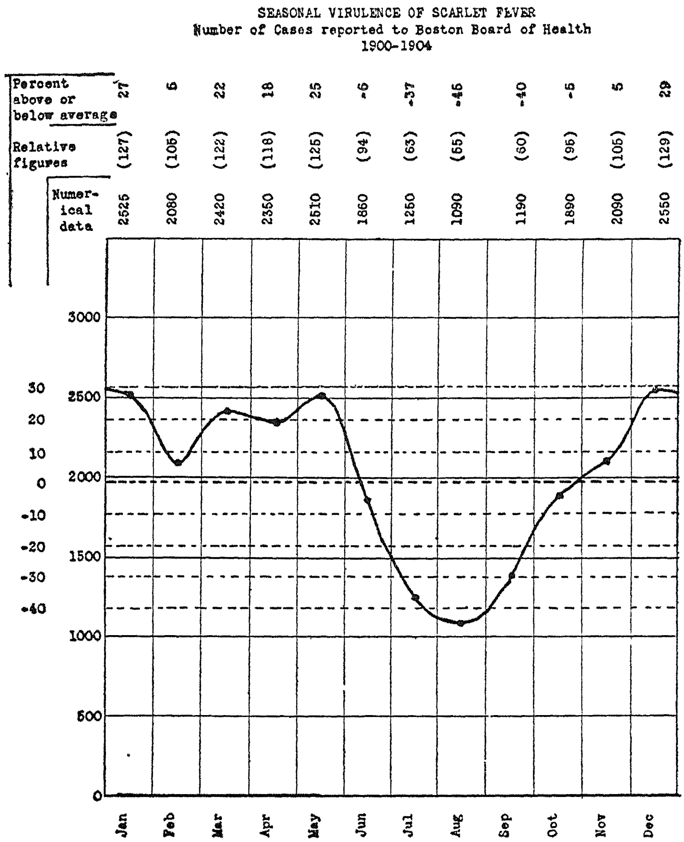


Fig. 218. Showing the Use of Relative (Percentage) Figures and a Rounded Curve.

Indeed, the statistician, and the chart-maker as well, would fail in one of his most obvious tasks, if he were to report as a significant rise or fall, a change which was wholly due to its cycles. Are we to conclude that the telephone business is disappearing because in a series showing hourly number of phone-calls, our last report is the number of phone-calls between twelve and one o'clock at midnight? It is true that between mid-afternoon and mid-night the telephone activity has dropped off almost entirely, but we must remember that it does this every night (with the possible exception of election-day) and that we deal here with a daily cycle. Are we to conclude that the cold-storage of eggs is a practise of the past, because our monthly report of warehouse stocks end with

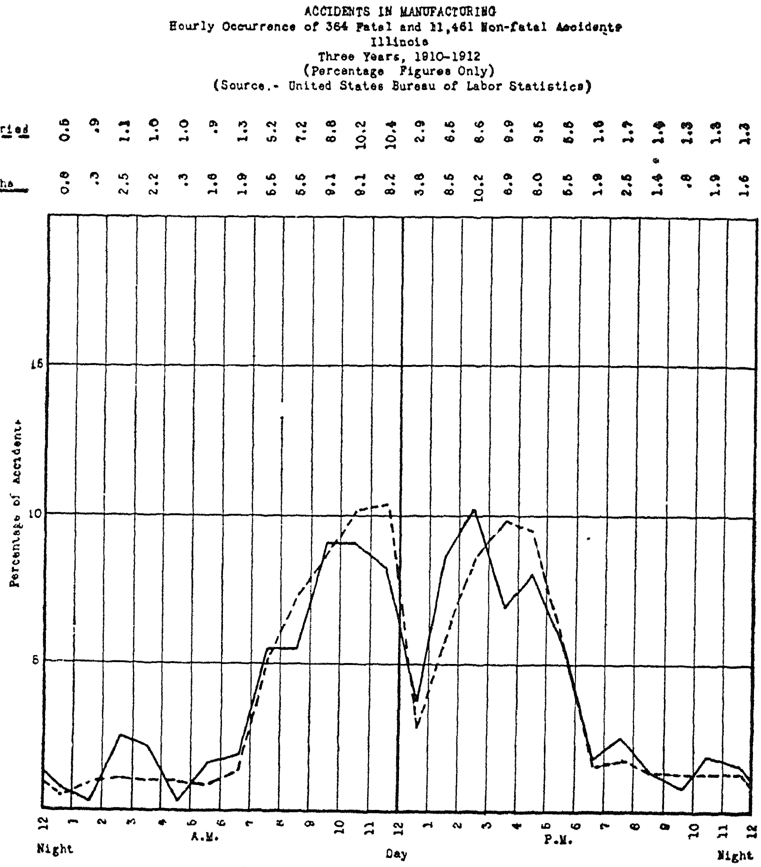


Fig. 219. Daily Cycles.

March, when as a matter of every-day knowledge there is an annual cycle and the stocks are always low at this time of the year?

The subject is more properly a statistical one than a charting one, but it is of such importance in the making of the specialized form of charts which follow that we will outline briefly some of the simpler methods in use. Our concern here is with the separation or elimination of the recurrent or cyclic fluctuations in an historical series. Ordinarily, in business, the seasonal fluctuation, that is, the annual cycle, is most important, and the following explanation will be limited to it. Other cycles may be similarly treated. The most elementary consideration in the analysis has been made obvious by the

COLD STORAGE HOLDINGS OF EGGS
Stocks of "Case Eggs" in Warehouses
United States
1916-1921
(Source:- Survey of Current Business)
(Monthly Average for Five Years, 1916-1920, = 100)

5 yr. Average 1916-1920	34	7	0	7	54	151	190	196	185	162	122	71
1920	42	9	1	3	58	139	183	186	173	144	104	49
1921	11	1	1	52	133	166	204	206	195	170	119	65

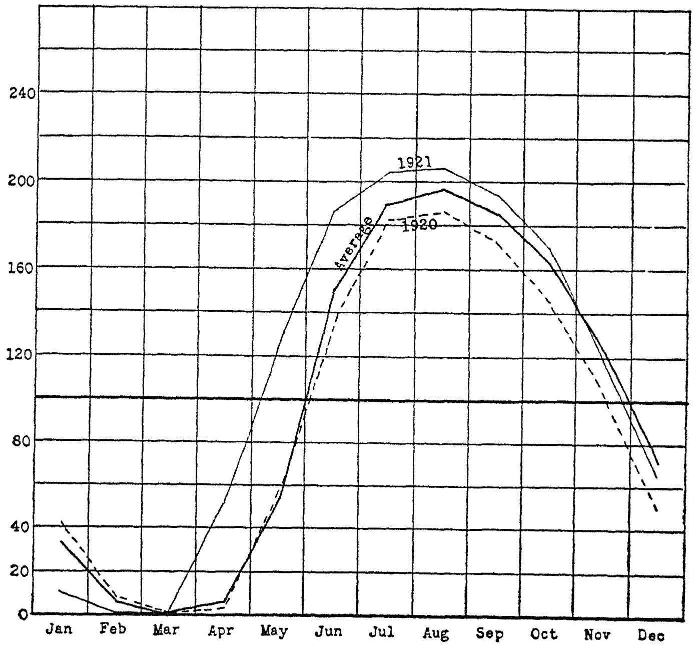


Fig. 220.

foregoing illustrations; namely that the relation between the last item in a series and the corresponding item in the previous cycle in the same series is of more importance than the relation between the last item and the immediately preceding item. In the last illustration, how do this year's October stocks compare with stocks of October last year, not how do this year's October stocks compare with this year's September stocks.

But every business man has progressed beyond this elementary stage. He asks to see the figures for the previous month in each of the last two cycles. For he knows that it is more important to see how the change in stocks from Septem-

ber to October this year compares with the change between the same months last year. We may generalize this by saying that

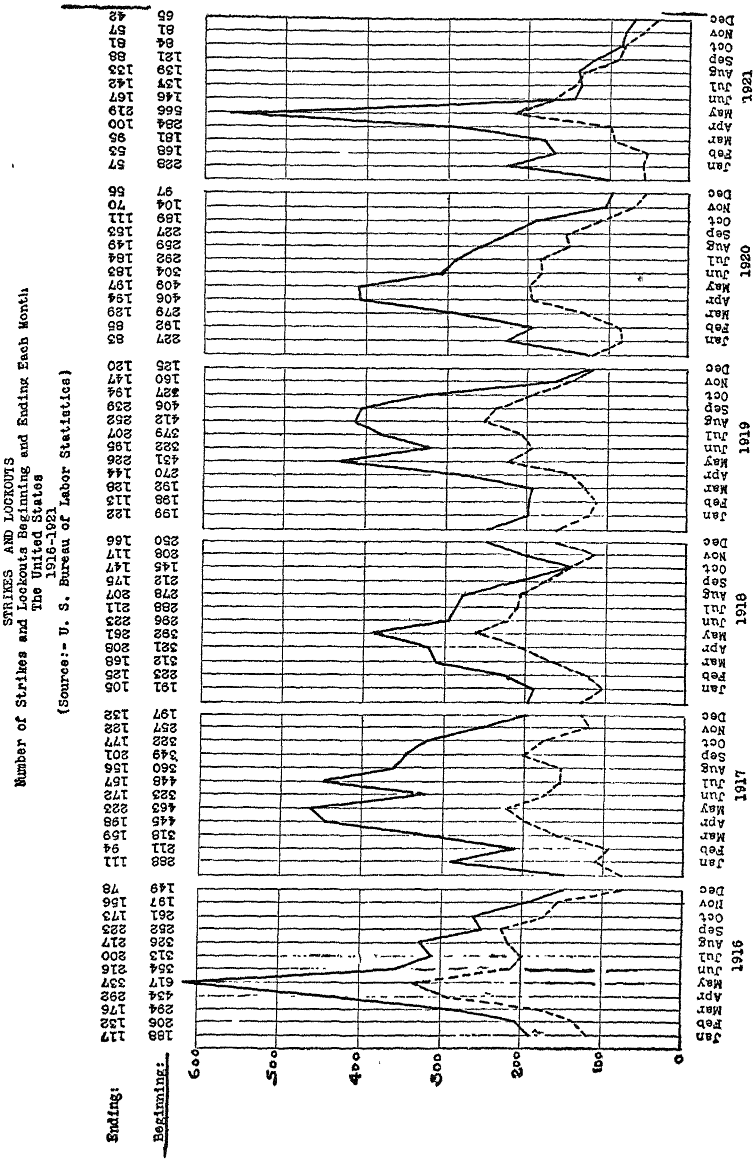


Fig. 221. Cycles May Change.

we are now concerned with the relation between the change in the last two items of the series and the corresponding change in

the corresponding two items in the previous cycle in the same series.[3] Now it is precisely this relation which the moving annual total, or average, described in the previous chapter, tells us. A little study will show that the moving total swal-

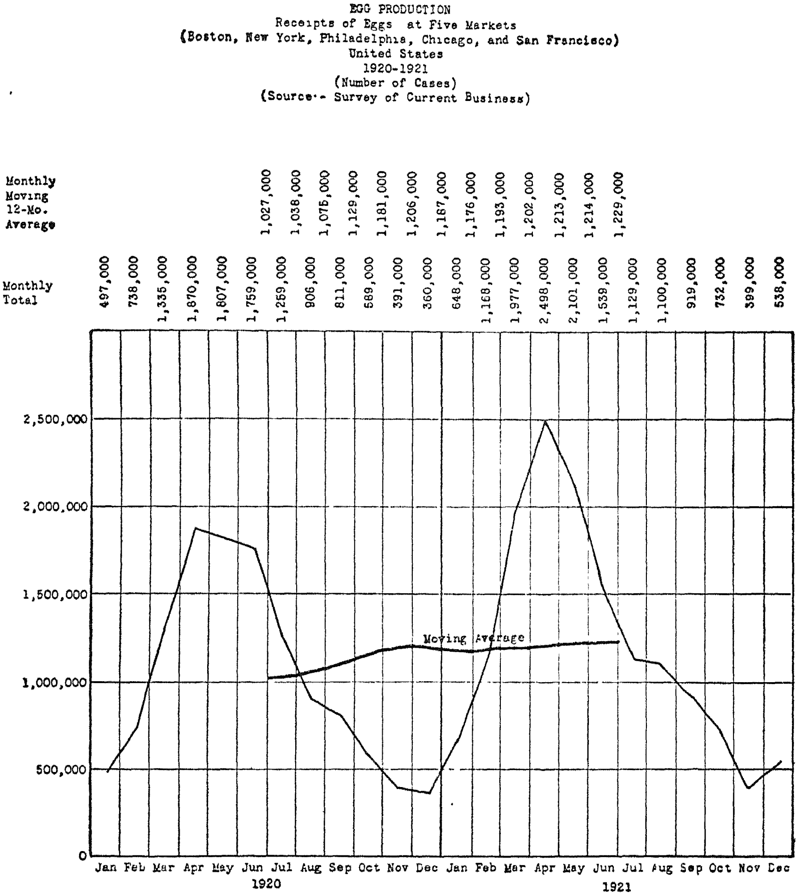


Fig. 222. The Moving Average Shows Trend.

lows up the cyclic variations by the simple process of swallowing up all the various items which make up the cycle. The

[3] As a matter of fact, the changes of the moving total from month to month are merely the differences between figures for the new month included and the old month excluded (one year previous). Hence, the work of computing a series of moving totals can progress largely at sight by comparing the month to be added with the month to be subtracted (one year earlier) and algebraicly adding this difference to the last moving total.

period covered by the moving total (or average) must of course be of the same length as the period of the cycle.[4] And the resulting changes in the moving total are merely the difference between changes in corresponding pairs of months in the two cycles. For this reason the moving annual total or average may be called the simplest method of eliminating seasonal (or cyclic) fluctuations, and determining the true secular (or long-time) movement.

While for many purposes, the trend as indicated by the moving total is a sufficient index of the nature of the more fundamental changes in the phenomenon, yet in the broad study of economic or trade movements, it still retains too many insignificant changes. It is true that the moving total smooths out all the periodically recurrent fluctuations. But it does not yet yield a simple series of perfectly regular change, that is, a straight line, or simple mathematical curve, which can be expressed by a mathematical equation or summarized in so simple a statement, as that, for example, "the population gains two per cent annually." In a much more precise sense the latter, that is a fitted straight line, parabolic curve, or other regular series, is called the "secular trend." It is also sometimes called the "normal" for the particular curve.

Fitting a straight line, or regular curve, to a historical series, is a matter of mathematical statistics into which we need not go, for it requires skill and judgment to which no simple rules of procedure apply.[5] It is sufficient to say that when this is taken into consideration the original series of data which we are analysing can be considered a combination of three elements, namely seasonal or cyclic fluctuations, a secular or normal trend, and secular fluctuations. And in the best statistical work both the former are often removed from the data before the curves are published. When this is done, the reader is advised of it by a simple statement to the effect that the figures published "are corrected for seasonal changes and normal growth." Fortunately he has no idea of the problems involved in this correction.

Accepting then, the moving total or average, as a satisfactory method of smoothing away all the insignificant and periodically

[4] When cycles are of varying lengths, this does not apply, for the moving total can only be made with uniform lengths or spans. The device next mentioned, however, the moving average, does not have this limitation and can be made co-extensive with the cycle.

[5] *Cf.* Chapter on Curve-Fitting.

recurrent fluctuations which often make monthly curves unsatisfactory—a means in short by which we can promptly plot the trend or general direction of underlying movements in an historical series—we turn to the question of determining the true nature of the cyclic, that is, the ascertaining of the true seasonal fluctuations. We wish now not to eliminate the cyclic changes in the data, but to eliminate everything else in the data and retain the cycle alone. How can we isolate the cycle? The simplest method and one which immediately suggests itself is to take a single cycle and forget the other cycles in the data. This gives us beyond peradventure the change within the

NEW INCORPORATIONS
Capital Invested in New Enterprises
Whose authorized capital equaled or exceeded $100,000
Principal States, U. S.
1921
(Source:- N. Y. Journal of Commerce)
(Figures in millions of dollars)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1,243	654	955	988	601	676	282	580	490	503	368	619	7,959
15.63	8.22	12.00	12.40	7.55	8.50	3.54	7.29	6.16	6.31	4.62	7.78	100.00

Fig. 223. The Seasonal Cycle Computed from One Cyclic Period.

cycle. If we wish, we can calculate the various months in this cycle as related to the total for the cycle, that is, change each month into a percentage of the total for the year, the latter being 100%.

The trouble with this crude use of a single year as an index or indicator of cyclic fluctuation is two-fold. For one thing it does not take any account of secular trend—which in the case of a young and rapidly growing business will be very marked—and as a result December sales may appear to be seasonally larger than January sales, though in fact they are really smaller, because every year the following January sales exceed the last December sales, just as within the calendar year the last December sales always exceeded the last January sales; the result of this error is to skew the seasonal fluctuation curve around in the cycle, tilting up one end of it, giving us a warped picture of the cyclic fluctuation, in which the warping or tilting may be so great as actually to shift the location of the peaks and valleys. If the data be the record of sales by an individual concrn, no matter how true a picture it may afford of the experience of the company, it does not give a true picture of the changes in the market, the seasonal variations in consumer demand.

The simplest way of correcting for the secular trend or general movement of the phenomenon is to take the months

NEW INCORPORATIONS
Capital Invested in New Enterprises
Those authorized capital equaled or exceeded $100,000
Principal States, U. S.
1921
(Source - N. Y. Journal of Commerce)
(Figures in millions of dollars)

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1921	Amount	1,243	654	955	988	601	676	282	580	490	503	368	619	7,959
	Moving Total	13,962	13,457	13,036	12,670	11,853	11,206	10,223	9,867	9,406	8,729	8,201	7,959	
	Percentage	8.92	4.86	7.33	7.81	5.07	6.03	2.75	5.88	5.21	5.76	4.49	7.78	71.89
	Corrected	12.25	6.77	10.19	10.87	7.06	8.39	3.93	8.19	7.26	8.02	6.25	10.82	100.00

Fig. 224. The Seasonal Computed from the Trend.

not as percentages of the total for the calendar or fiscal year of fixed span, but to take them as percentages of the moving total ("for the current twelve months"), for the same months. The results will no longer add up to 100%, but will be less than 100% if the moving total has fallen and more if it has risen. The monthly percentages of the moving totals must therefore be summed up and corrected so that their sum equals 100% (by dividing them by their sum). The result of this process may be taken as in most cases an entirely satisfactory record of the typical seasonal fluctuation, during one year.

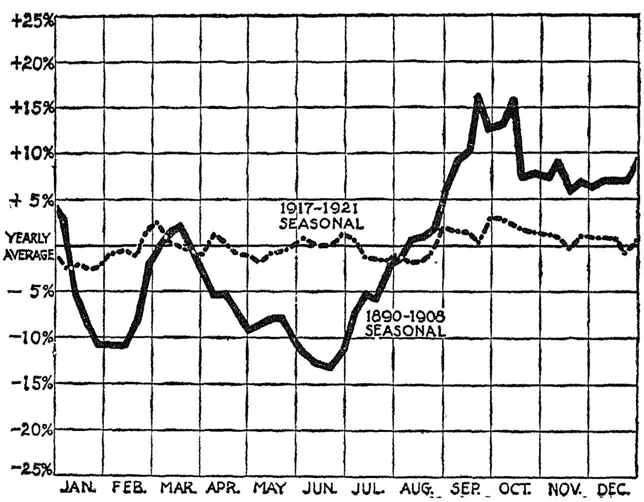


Fig. 225. A Remarkable Case of Changing Cycle Fluctuations.

Typical seasonal changes in interest rates on 60 to 90 day commercial paper for the years 1890 to 1908 and the years 1917 to 1921. Weekly variations are shown as percentage deviations from the annual average.—*Permission of Mr. Carl Snyder.*

The second objection to the method still remains, however. This objection is that the cycle is estimated upon a single year's experience only. The cycle shown by another year might be somewhat different. And how do we know that one year is any more representative of actual conditions than another. Of course, in the case of businesses (or other phenomena) effected by the war (and this includes most business and economic records), we might quickly throw out the wartime record, as being wholly unreliable. But in the absence of special reasons for discarding certain periods or records as unrepresentative, we may be confronted with many years of equal significance which yield different cyclic curves. And in such cases it would be wrong to trust one entirely and discriminate against the rest. The obvious thing to do is to calculate the seasonals for each of these years, by the method above described, and then average them together, to get an average seasonal. The resulting curve would meet the second objection and be representative of the entire experience, for which records are available.

Perhaps the most important use of seasonals in ordinary business statistics, is the calculation of "quotas," or planned "schedules" for the future. In a sales department, for example, the quotas assigned in advance to the salesman, or to the sales districts, should be as fair as possible, and to assure this, the typical seasonal fluctuations should be known. Our problem then becomes slightly different. We no longer want the seasonals most typical of the entire experience of the past, but we want the seasonals which may be considered most typical of the immediate future. A simple average of the seasonals for many years past would give too little importance, perhaps, to recent developments. It may be that, through advertising, or through changes in market conditions, the consumer demand has been shifted about in the year (usually to become more level, that is, regular). For such developments it is plain that the later years are more truly representative than the earlier ones.

In the calculation of quotas, therefore, it is well to "weight" the later years more heavily than the earlier ones before averaging. Ordinarily it is satisfactory to weight each year twice as heavily as the preceding year. Other weighting systems can be used, but this particular arrangement leads to a most easily calculated average seasonal which has been devised and

used by the author for a long time under the convenient, though somewhat loose, name of the "compounded average."[6] It has the advantage of being easily carried on from year to year without extensive re-calculations, an important factor in a busy office, and it also avoids all question of how many back years to include, by making all except the last four or five

NEW INCORPORATIONS
Capital Invested in New Enterprises
Those authorized capital equaled or exceeded $100,000
Principal States, U. S.
1918-1921
(Source - N. Y. Journal of Commerce)
(Figures in millions of dollars)

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1918	Percentage	6.28	4.12	4.69	6.35	8.05	6.23	5.57	4.77	7.24	4.91	5.27	5.42	68.90
	Corrected	9.12	5.98	6.80	9.21	11.70	9.04	8.08	6.92	10.51	7.13	7.64	7.87	100.00
1919	Percentage	18.87	11.79	12.71	16.22	20.68	27.05	24.15	12.53	23.49	22.44	11.43	8.52	209.88
	Corrected	9.00	5.62	6.08	7.74	9.86	12.80	11.50	5.98	11.20	10.71	5.45	4.06	100.00
Total	1918 and 1919	18.12	11.60	12.88	16.95	21.56	21.84	19.58	12.90	21.71	17.84	13.09	11.93	200.00
	Average	9.06	5.80	6.44	8.46	10.78	10.92	9.79	6.45	10.81	8.92	6.54	5.97	100.00
1920	Percentage	15.76	7.57	8.44	7.90	7.97	7.41	7.11	5.27	5.64	7.54	5.88	5.75	92.24
	Corrected	17.10	8.21	9.16	8.57	8.62	8.04	7.71	5.71	6.11	8.17	6.37	6.23	100.00
Total	1920 and Average	26.16	14.01	15.60	17.03	19.40	18.96	17.50	12.16	16.92	17.11	12.91	12.20	200.00
	Compound Average	13.08	7.00	7.80	8.52	9.70	9.48	8.75	6.08	8.46	8.55	6.46	6.10	100.00
1921	Percentage	8.92	4.86	7.33	7.81	5.07	6.03	2.75	5.88	5.21	5.76	4.49	7.78	71.89
	Corrected	12.25	6.77	10.19	10.87	7.06	8.39	3.93	8.19	7.26	8.02	6.25	10.82	100.00
Total	1921 and Average	25.33	13.77	17.99	19.39	16.76	17.87	12.68	14.27	15.72	16.57	12.71	16.92	200.00
	Compound Average	12.66	6.89	9.00	9.69	8.38	8.93	6.34	7.14	7.86	8.29	6.35	8.46	100.00

Fig. 226. The "Compounded Average" Seasonal.

negligible. The trick is to average the seasonals for the first two years (which can be done at sight), and than average the resulting average with the next year to get a new average, continuing this process through the years and always working by inspection.

With the method here outlined to use when you wish to ascertain the seasonal fluctuations in your data, using the

[6] Of course, in this "compounded average" the weighting is not two to one for the first two years, but with the exception of the first year the weighting is in this ratio throughout, and in a very few years the importance of the first year is rendered so negligible as to be lost.

By other weighting systems, it is meant that ratios of three to one, or of one to two-thirds, or of one to three-fourths, or the like, can also be easily used and currently maintained (that is, brought up to date) almost by inspection.

The theory of the compounded average is very simple, and appears to be sounder than that of any fixed average seasonal. It is believed to be an original contribution to the science of averages, which should have particular value in economic work with phenomena undergoing changes in seasonal fluctuations. A very spectacular case of such a phenomenon was the behavior of the interest-rates for loans in New York after the establishment of the Federal reserve system, when a previously marked seasonal was almost entirely wiped out in a few years. The compounded average affords a sort of moving or progressive seasonal well adapted to such cases. And, in the ease with which it is brought up to date, it is, mechanistically, a decided labor-saver and time-saver.

"compounded average" in the place of the simple average for quota-making, and with the moving total[7] and average previously described for the elimination of the seasonal when you wish the real underlying movement, loosely called the secular trend, in your data, you are equipped with the mathematical means necessary for the successful use of the following charts. Apart from the need of the cyclic change in quota-making, the usual need is for the secular trend and the latter is indeed useful not merely for the following charts, but for a wide variety of purposes in statistical work. It is therefore the more important trick to have up your sleeve, in attacking either business or sociological statistics. However far it may fall short of a true secular trend, it still gives a significant and easily understood smoothed curve. Though still little known to the average executive, it is proving extremely popular among those who use it, and has been credited by some business statisticians, chiefly those who use the device described in the next chapter, as being the only part of the data in which the executive should be interested.

[7] By an oversight, the tables in the discussion of the compounded average all show the months as percentages of the moving total for "previous" 12 months. It is obvious that the moving total of "current" 12 months should have been used.

Chapter XXII

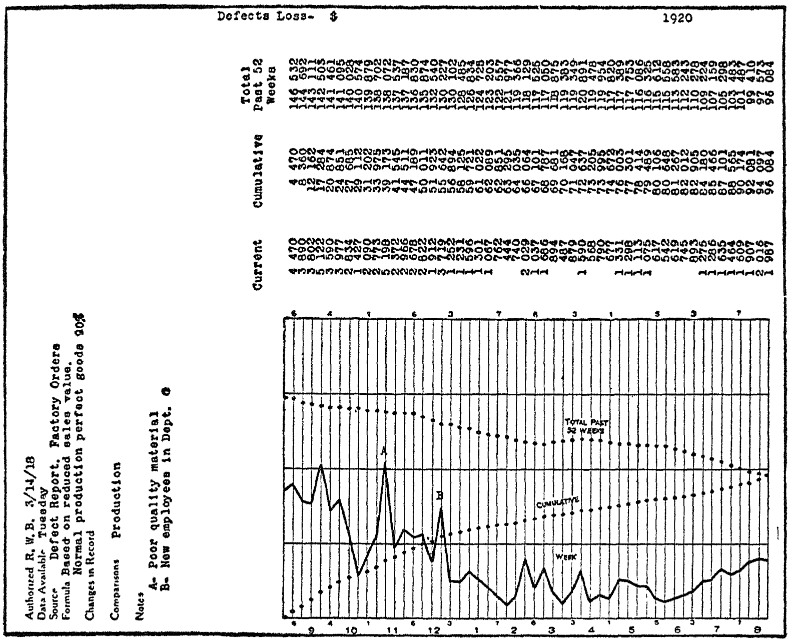
ZEE-CHARTS

In this chapter we enter the accountant's paradise, and instead of simplifying our data and presentation, we multiply it three-fold, by adding to each original series of data, its cumulative and its moving total series. The result is a chart which shows simply and coherently everything about the data which can be shown. It takes much space, for each important period (i.e. year or month) of data should be given a separate chart, and its use is therefore better restricted to a few series, whose importance is sufficient to justify their treatment in this thorough and painstaking way. In its way, this chart is the last word in the analysis of and research into past historical data.[1]

The "Zee-chart" gained its name from the fact that its three curves roughly form the letter "Z." These three curves are, first, the curve of the original data, second, the cumulative curve, and third, the moving-total curve. In common practice there are three kinds of these charts, depending upon the time period of the original data. Where the original data consists of monthly figures, the chart shows twelve of these figures to form one year; when the data is weekly, fifty-two weekly figures are combined in one chart to show one year; and when the data is daily, thirty or thirty-one days are combined to show in one chart a month. In the first two cases the moving-total is an annual one and in the last case it is a

[1] The Zee-chart is rumored to be of German origin, but appears to have had a somewhat later independent American discovery. It has received its greatest development at the hands of Mr. Willard C. Brinton, consulting engineer and author of *Graphic Methods for Presenting Facts*. The present writer is informed that the combination of monthly and moving total curves on standard scale combinations was worked out by Mr. T. R. Robinson, and the addition of the cumulative curve was suggested by Mr. Wallace Clark. Of late, the Zee-chart has been further modified in its form by Mr. Arthur R. Burnet who has also suggested the omission of scale-figures to focus attention upon the curves, and has invented scale-finding machinery to facilitate plotting.

monthly one, though it is to be noted that in most businesses a monthly moving-total has little significance. The Zee-charts can, however, be made up of any combination of time units desired. The most practicable one, and the one which will be herein described, is the Zee-chart of monthly data, twelve months comprising one chart. Its ready adaptability to the needs of the business man and accountant makes it extremely useful for recording data in these fields.



Permission of the Ronald Press.

Fig. 227. A Year by Weeks.

Showing the arrangement and methods advocated by Mr. Burnet.

By this time, if you have understood the last two chapters, you will be protesting that it is not practicable to show a moving total curve on the same chart with the curve of the original series. Why? Because the moving total for twelve months is twelve times as great as the average monthly items, and if the moving total curve is to be shown, the curve of monthly figures will lie very close to the bottom of the chart,

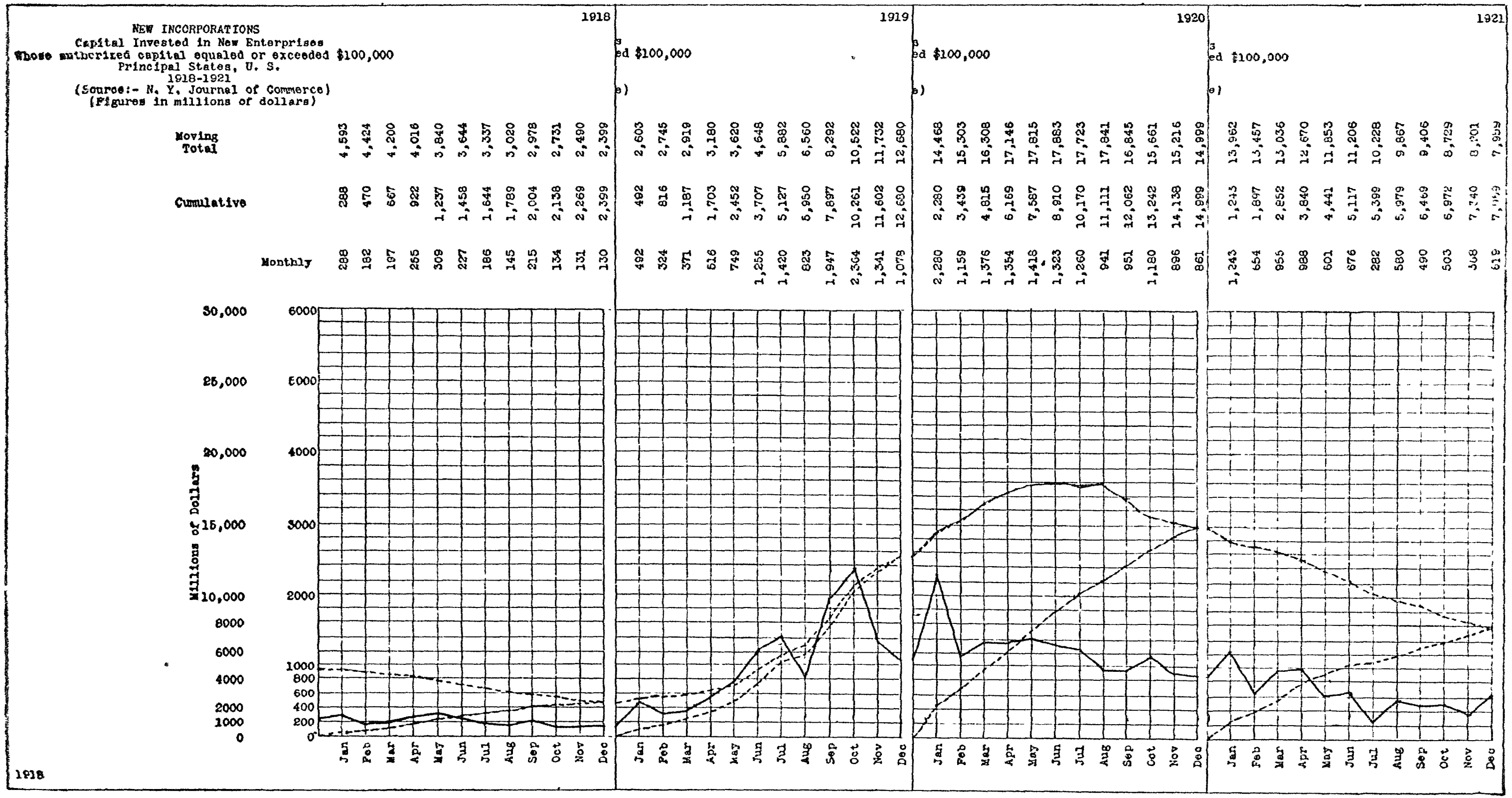


Fig. 228. Four Zee-Charts Forming a Single Series.

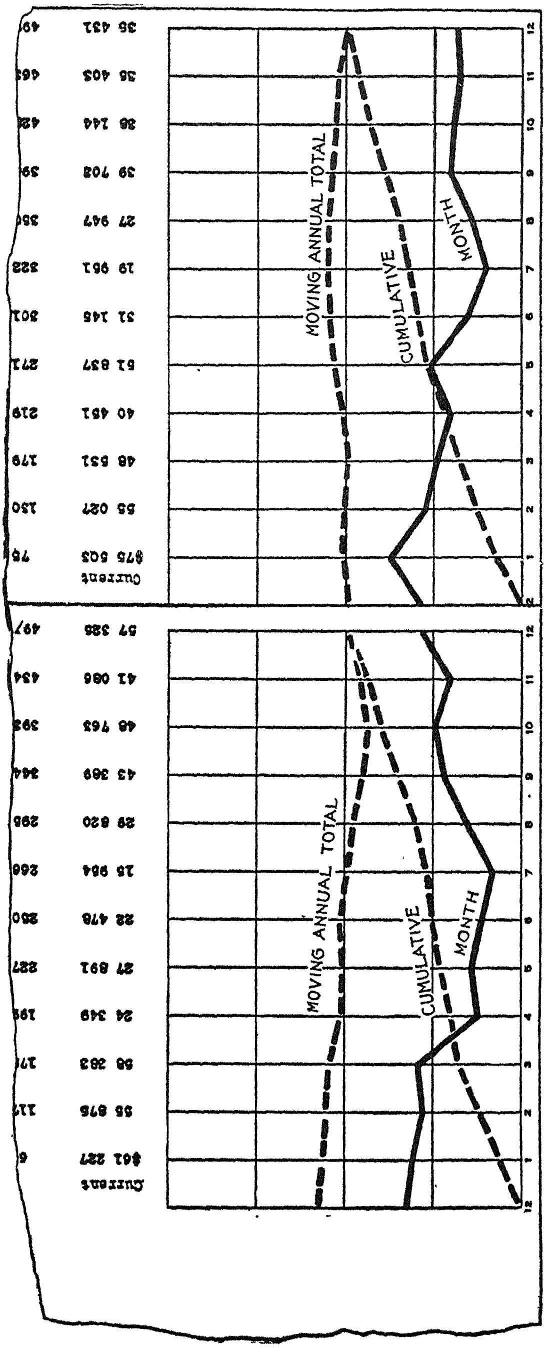
Notice the different scales for monthly and annual data and the corresponding positions of the data captions. In the originals the annual moving total and cumulative data and curves and scale are in red.

its fluctuations hardly visible. The objection is well taken, but the Zee-chart cleverly dodges this difficulty. How? By the simple device of using two scales. A large scale is used for the original monthly figures, but a small scale is used for the cumulative and moving total curves, whereby they are brought down on the chart to not more than two or three times the height of the monthly curve. The ratio between these two scales has been more or less standardized, the monthly-curve scale being five times as great as the annual cumulative and moving total scale. When the data are weekly and the moving total is a 52-weeks total, the ratio is larger, the weekly scale being twenty times as great as the cumulative and moving total scale. When the data is daily and the moving total monthly, the ratio is again different, the daily scale being ten times as great as the monthly cumulative and moving total scale.[2]

Standardized practice also has it that a color distinction should be observed between the two scales and their corresponding curves and data. You should enter the individual series in the data at the top of the chart in black, plot the curve therefore in black and enter the scale itself in black lettering. But the cumulative and moving-total figures should be entered in red in the data at the top of the chart, likewise their curves should be drawn in red ink and their common scale entered in red letters. This distinction is an excellent one as it causes the curve of original data to stand out most prominently while the moving total curve or trend is perfectly clear, some distance higher on the chart.

As you have seen in an earlier chapter, the moving total of a series can be plotted at any point within its period, and can be a moving total either of the current twelve months, of the following twelve months, or of the past twelve months. It is the last kind of moving total alone which is used in the Zee-charts. From this fact, two advantages arise. In the first place all three curves are entered at the same time on the same ordinate and a hasty reader who has no time to analyze the figures, gains no false impression that the chart is not

[2] A very shrewd suggestion has been made by Mr. Arthur R. Burnet, consulting statistician and graphic expert, that the scale-figures be omitted from charts which are to be shown to the non-technical business man, in order that his attention may not be diverted from the behavior of the curves. The fact that data is always included in the Zee-chart, and hence scales can always be ascertained though not calibrated, gives to this suggestion unusual merit.

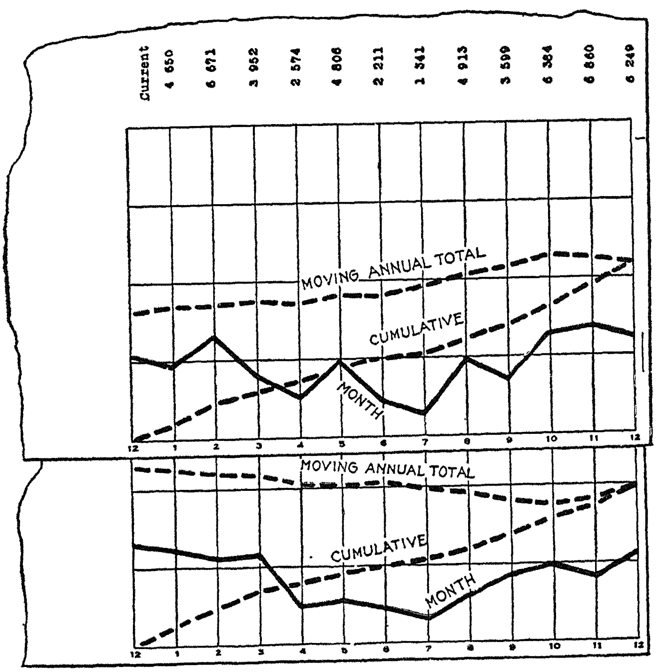


Permission of Mr. Arthur R. Burnet and the Ronald Press,

Fig. 229. Two Charts Fanned Out For Successive Years.

thoroughly up-to-date, as he is likely to when the moving total is plotted upon an earlier ordinate. In the second place, the cumulative and moving total curves always come together at the end of each chart by this device, since it is obvious that the cumulative for twelve months is the same as the total for the past twelve months. This coincidence of the two curves is an important element in the simplicity of the chart and also automatically checks both computing and plotting.

The Zee-chart should be prepared upon the same type of paper as other historical curves, that is, its field should be positioned in the lower righthand corner of the page, close to the edge of the paper. For Zee-charts are generally prepared in sets for data extending over many years and requiring as many charts in the set as there are years. It is important, therefore, that the reader be able to fan the charts out, each overlapping the next so as to afford a general view of the entire series of years in a small space. The narrow margin to the right of the chart on each sheet produces gaps between the



Permission of Mr. Arthur R. Burnet and the Ronald Press.

Fig. 230. Two Charts Fanned Upward to Study Seasonals.

charts when laid out in this way but these gaps are of benefit, forming slight breaks between years without destroying the general continuity of the set of charts. The low position of the field upon the paper facilitates the study of seasonal fluctuations by overlapping charts one above the other. Needless to say, the scale on all charts belonging to a set, should be uniform and rigidly maintained within the set in order that the curve running through the charts may be homogeneous.

The scale on a Zee-chart should be somewhat smaller than on most historical curves, for the reason that you probably expect to continue the Zee-chart series in the future, and you must allow ample margin for future growth of the business and future peaks which will rise higher than the past peaks. To find the scale for a Zee-chart, therefore, look back over the series of moving-total figures in your data and locate the highest peak figure in the past, and then select a scale which

Observed Peak in Entire Moving Total Series	Type of Field to use for the Chart	CURRENT MONTHLY SERIES				CUMULATIVE AND MOVING TOTAL SERIES			
		Scale-figure first Abscissa	Value of Data to Plot	Rule to use in Plotting	Plotting Value of Rule-unit	Scale-figure first Abscissa	Value of Data to Plot	Rule to use in Plotting	Plotting Value of Rule-unit
100 - 120	40	2	N/2	40	1	10	N	40	10
120 - 150	50	2	N	10	10	10	N	50	10
150 - 180	60	2	N/2	60	1	10	N	60	10
160 - 240	40	4	N/4	40	1	20	N/2	40	10
240 - 300	40	5	N	20	10	25	N	10	100
300 - 360	60	4	N/4	60	1	20	N/2	60	10
360 - 450	60	5	N	30	10	25	2N	30	100
450 - 600	40	10	N	40	10	50	N	20	100
600 - 750	50	10	N	50	10	50	2N	50	100
750 - 900	60	10	N	60	10	50	N	30	100
900 -1000	40	20	N/2	40	10	100	N	40	100

Fig. 231. Table for Zee-chart Scales.
For standard chart-fields six inches high.

will bring this figure about half way up the chart. Selection of the scale is perhaps one of the most difficult parts of Zee-chart making, and it is convenient to use a table of scales similar to the table for historical scales given in a previous chapter, modified only by increasing the key numbers one-third to lower the peaks to half way up the chart.

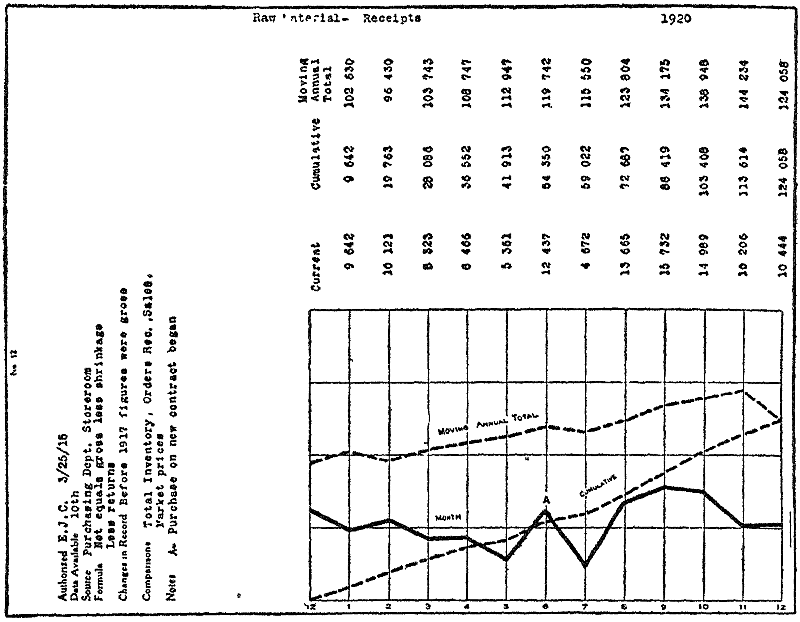
One of the most interesting features of Zee-charts is their use in what is called "light analysis." Light analysis is a method of comparing two curves or charts drawn on similar scales. You place the two charts together, one upon the

other, and hold both up to the light. Both curves will then be visible, enabling you to compare them minutely. In some offices where a large number of Zee-charts are used, a machine called a "light box" is kept for this work. The light box consists of an electric light underneath a piece of ground glass, which throws a uniform illumination up to the paper laid upon it. The charts should therefore be made on highly translucent[3] paper, and the field should be positioned with absolute uniformity on every sheet. These are considerations which hold not merely for Zee-charts but for all curve-charts to be subjected to light analysis.

A number of attempts have been made to adopt the Zee-chart to current operation control. It is easily seen that by extending or projecting the cumulative curve on any up-to-date chart in which the end of its period (e.g. year) has not yet been reached, an estimate can be quickly made of the probable amount to which the future months will bring the entire year's total. Moreover, if a quota for the year has been made in advance, the cumulative for this quota can be easily plotted in pencil or, best of all, in yellow ink and a comparison between the red cumulative for actual performance with the yellow quota cumulative will show how well or poorly the quota is being fulfilled. A slightly different method is that of showing the quota cumulative as a straight sloping line, and entering the curve figures and cumulative as percentages of the amount which would have been necessary to meet this quota. But these various methods have not been so successful as to find general acceptance, as they tend to fill up the chart with too much detail.

The plain fact is that the Zee-chart is a "looking-backward" chart. The best that can be said of it is that it is an ideal method of historical research. The sales of your chief line or department should be plotted in this way, in order that you may study their past history to the best advantage. The chart is designed to show you at once the individual monthly or periodic fluctuation, their general trend or moving total, and the cumulative or total to date for each year, and from these three accounting elements for each year, you can see just when sales began to fall off or rise, what the seasonal fluctuation was, and how each year's individual progress compared with

[3] The paper may be highly translucent without being in the least transparent.



Permission of the Ronald Press.

Fig. 232. Mr. Burnet's Arrangement.

that of other years. The Zee-chart is, among amount-of-change charts, the last word in historical research and detail.[4]

[4] An excellent description of the Zee-chart is to be found in a series of articles by Mr. Arthur R. Burnet in *Management Engineering*, beginning Sept. 1921, pp. 153–160.

The first American description is that by Mr. John Wenzel in an early issue of the *Scientific American Magazine*.

The best adaptation for forecasting and quota-measuring has been made by Mr. John Scoville, Maxwell Motors, Detroit.

Chapter XXIII

PROGRESS CHARTS

From the paradise of the accountant and historian let us step into the paradise of the operating executive. The operating executive is the man who sees to it that things are done. His interest begins and ends with the job to be done, how much of it has actually been done, and how much remains to be done. When told that a job has not yet been finished, he is not interested in excuses and reasons, he is not interested in why or when the performance fell below the schedule, he is only interested in the amount remaining to be done and the job of getting it done. We shall not expect to find him satisfied with a retrospective or "looking-backward" chart. For him, the "looking-forward" chart!

Far and away the most "forward-looking" chart known is the "progress" chart. It is the product of the man who was probably the greatest engineer America has ever produced, the late Mr. H. L. Gantt.[1] It is significant that this chart method was devised by an engineering type of mind, for it is admirably adapted to an executive control of operation. Compared with its dynamic influence on the actual control of operations, nearly all other types of charts seem to justify the assumption that the word "statistical" is derived from the word "static."

---

[1] Henry Laurence Gantt was born in Maryland on May 20, 1861. He received the degree of bachelor of arts in 1880 from Johns Hopkins University and in 1884 the degree of mechanical engineer from Stevens Institute of Technology. He died November 23, 1919, at his home in Montclair, N. J.

Mr. Gantt was associated with Frederick W. Taylor in his early work at the Midvale and Bethlehem Steel Companies and a few years later established his own consulting practice as an industrial engineer, which he carried on until his death.

Among his clients were many of the most advanced manufacturing companies of this country. During the war he devoted his entire time as well as that of his staff to the solution of the Government's problems of production and management. He acted in a consulting capacity for the Ordnance Department, the Naval Aircraft, the Shipping Board, and the Emergency Fleet Corporation.

Mr. Gantt developed a method of paying workmen according to the results they

The Gantt progress-chart presupposes a definite detailed schedule or plan made out in advance and generally called the "quota." The chart itself merely measures the subsequent actual performance, when it takes place, against this predetermined schedule or quota, and shows emphatically whether or not this quota is being met. The chart shows incidentally how much of each past month's quota has been accomplished, but primarily it shows how much of the cumulative or total quota to date has been accomplished. The chart never shows trend or moving total, nor does it necessarily show even the individual monthly figures; its main function is to show how much of the schedule has been performed up-to-date and how much remains to be done.

As we might have expected in a chart for an operating executive, the progress-chart compresses its information into very small space. Where the Zee-chart expanded every series of figures three-fold and used a separate sheet of paper for each series, the progress-chart combines twenty, thirty, or even more, series upon a single page. An entire business or industry can be summarized upon one of these charts, each of its thirty or more items being in turn shown in detail with thirty or more sub-divisions on subordinate charts. In the course of time, the Gantt progress-chart will come to be recognized as the *sine qua non* of management, whether it be sales management, office management, or production and factory management.

Strictly speaking, the Gantt progress-chart is not a curve-chart at all. It is rather a horizontal bar-chart, very peculiarly constructed. But the relation between bar-charts and amount-of-change curve-charts is so intimate that it can best be examined here. If you like to so consider it, the Gantt progress-chart is a combination of many curve-charts, each flattened out into one dimension and all placed close together

---

accomplished, which is known as the Gantt Task and Bonus; he developed the theory that the cost of an article includes only those expenses actually incurred in the production of that article, and that the expense of maintaining one machine in idleness can not be charged into the cost of the output of another machine. In accordance with this theory, he worked out a method of arriving at costs of idleness and of work; he originated the Gantt Chart, which compares the amount of work done in a given time with what *should* be done and emphasizes the reasons for failure to attain that standard; he introduced a change in the installation of management methods from the old type, which organized from the top down, to a new type which builds from the bottom up. "Work, Wages and Profits" (1910), "Industrial Leadership" (1916), and "Organizing for Work" (1919), are the titles of the most important books written by Mr. Gantt.—*Wallace Clark.*

on a single chart. Or if you prefer, the Gantt progress-chart is a series of horizontal bar-charts with cumulative bar-charts superimposed.

The scale of the progress-chart is one of its most interesting features. At first glance, there appear to be as many scales on the progress-chart as there are bars, or items. That is to say, every bar on the progress-chart appears to have its own scale or set of values for the horizontal distances through which it passes. And, unlike all other charts, the horizontal distances appear to have no equal and uniform values throughout the length even of a single scale; on the contrary, the values given to equal horizontal distances, or spaces between vertical lines, seem to change weirdly from space to space throughout each individual line. But the secret of the puzzle is very simple. The common and proper scale for progress-charts is "time." Each equal distance represents an equal unit of time, and you will find a time scale placed at the top of each progress-chart, one single uniform scale for the entire chart. Time, then, is the measure or unit of measurement against which performance is measured. In Mr. Gantt's words, "time is the one common thread running through all operations," and time is therefore the one basis on which all performance should be judged by executives.

Now it is obvious that the number of dollars worth of goods sold during a month cannot be taken directly as a part of, or measured directly in terms of, the number of days in the month. The amount of sales, production, or other performance, and the length of time involved in the performance are numerically incommensurable quantities. We must therefore have a ratio or co-efficient between the two, that is, we must make up our minds that a certain unit of time is to equal a certain volume of sales or other performance. Then the actual amount of performance can be judged in terms of this predetermined quantity, which has been decided upon for the given period of time. And these predetermined quantities are the items which at first glance appeared to form the irregular scales for each bar.

In short, the progress-chart measures actual performance in terms of a standard. This standard is shown by the small figures in each space and is graphically represented by the entire space itself. Actual accomplishment is graphically recorded by a bar drawn across the part of the space which

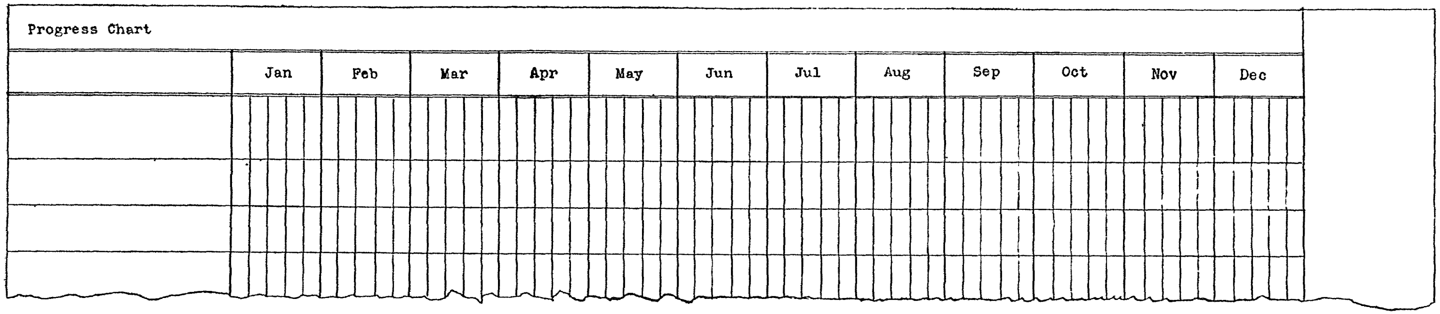


Fig. 233. Detail of a Progress Chart Form—Blank.

This sample carries five spaces to the time-period and twelve periods to the page, and is suitable for monthly data. The five spaces per period are convenient for percentage subdivisions. Any other combinations of time periods and subdivisions can be used. The months have been entered on this sheet, as well as a title; otherwise it is a blank. The full page size is 11 x 17 inches, binding on the right-hand edge.

corresponds to the percentage (of the standard), which has been accomplished. In other words, the standard for each space (or period of time) is considered to be 100% for that period of time, and the actual accomplishment during that period of time is taken as percentage of this standard, and shown by a 100% bar, in which the shaded portion represents the part accomplished, the unshaded portion the part not accomplished. This is literally true of the light lines or bars which begin afresh at the beginning of each new space and indicate the monthly or periodic performance or accomplishment. The method is not quite so simple, however, in the case of the cumulative performance or accomplishment, that is, the accomplishment from the beginning of the entire period shown on the chart, to date. This cumulative of performance is shown by heavy bars. Here (1) all the standard cumulatives which are less than the accomplishment cumulatives are considered wholly performed and 100% done, and the heavy cumulative bar is drawn entirely across them; (2) the remainder of the accomplishment cumulative, after subtracting the last standard cumulative, is taken as a percentage of the next individual period standard and (3) the cumulative bar is drawn correspondingly across the corresponding percentage of the last period space.

A simple example will make this clear. Suppose we are allowed by the publishers of this book ten months in which

PROPOSED SCHEDULE

Month	Quota
1	3
2	1
3	4
4	6
5	6
6	6
7	6
8	6
9	6
10	6
Total	50

Fig. 234.

to prepare it. As a matter of fact, the book is the result of many years of study—but that is another story. Being a

Progress Chart	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Dictation	3 3	1 4	4 8	6 14	6 20	6 26	6 32	6 38	6 44	6 50	0 50	0 50

Fig. 235. The Quota is Entered For a Year Ahead.

Here is the form, ready for the graphic record to be entered monthly—the schedule or quota being determined in advance for the entire period. In each space the monthly quota is at the left side and the quota cumulative at the right side of the space.

methodical sort of person, we sit down and prepare an outline of the book, and discover that it will take about fifty chapters in which to tell all that you should know about the subject. We then prepare a schedule showing how long it will take us to write each chapter, and decide that we can write the first three chapetrs in the first month, one chapter in the second, then four and thereafter six chapters a month. Next we prepare a progress-chart of this work showing ten months on the chart, one space for each month. We enter the number of chapters to be done each month in the upper left hand corner of the space for each month. We also enter the cumulative in the upper right hand corners of each space, showing that at the end of the second month we will have written four chapters, by the third month eight chapters, by the fourth, fourteen, and so on, until at the end of the tenth month fifty chapters are written. This checks our monthly schedule.

Time passes and we are writing, patient reader—though you might not guess it—we are writing with meticulous and painstaking care. If at the end of the first month, only one and one-

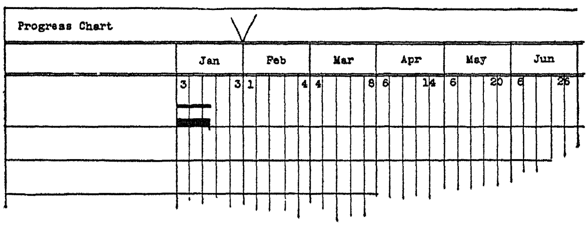


Fig. 236. The Chart on Jan. 31st.
The light line records a performance of 50% of the month's quota; the heavy line records 50% of the first month's quota cumulative. The V-shaped mark shows date of last entry.

half chapters have been completed halfway across the first space, when we should have written three chapters, we draw two lines, one light and one heavy, the light one being above the heavy one. By this we know that at the end of the first month we have only done 50% of the month's quota. In the second month we finish the second and also a third chapter. A new, light or monthly bar can be drawn all the way across the second space and a second light bar halfway across, above it, showing that we had done our bit and 50% more in the second month. But as we were short in the first month's work we are still short to date, having done three chapters

when we should have done four. These three chapters finish our quota for the first month only and leave nothing accomplished out of the second month's quota. We therefore draw the heavy cumulative bar completely across the first space

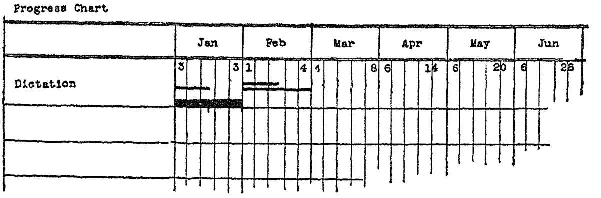


Fig. 237. The Chart on Feb. 28th.
The light lines show 150% of the second month's quota performed in the second month; the heavy bar shows total performance to date (Feb. 28) just one month behind schedule.

but not into the second space at all. A glance at the chart now shows us that we are short one month's work (though we have exceeded our quota in the second month). If in the third month, in an excess of energy, we write six more chapters, we shall have done 150% of the third month's quota of four chapters and shall therefore draw a light line all the way across the space, and over it another light line one half of the way across the space. But the cumulative or total to date will be nine, enabling us to draw the heavy bar all the way across the page to the end of the third month and one-sixth of the way across the fourth space, since the fourth months' quota was six chapters, of which we have done one. A glance at the chart now shows us that we are ahead of schedule. This illustration may seem wholly personal, but the sales or factory manager who cannot see in it a method applicable to his own problems is not worth his salt.

You will now understand something of the unique merit of the progress-chart. It has an uncanny power of making human judgment. It does not merely record what has been done, but in addition thereto it records whether this accomplishment has been good, bad, or indifferent. It does not merely put the question, but it also gives the answer. It weighs every fact in the balance and states in unmistakable terms the judgment. If you have fallen down on your job, the chart does not waste emphasis on why, when, or how you fell down, but places before you in a way which you cannot dodge the fact that you have fallen down. It is of course

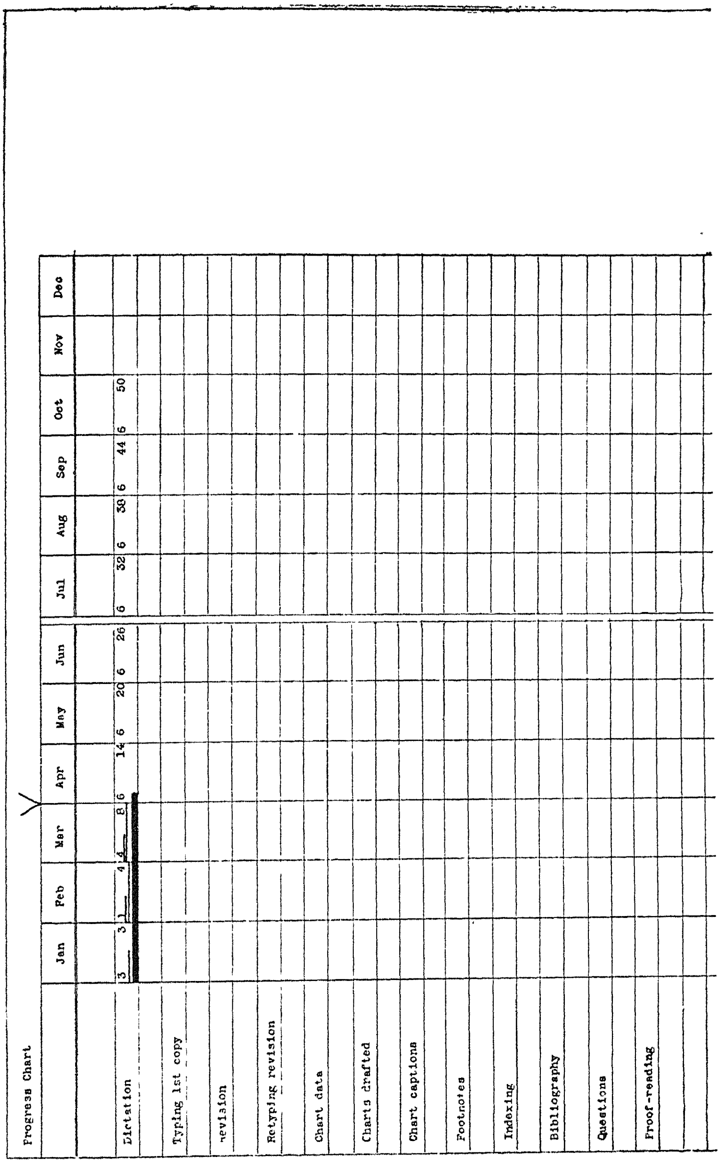
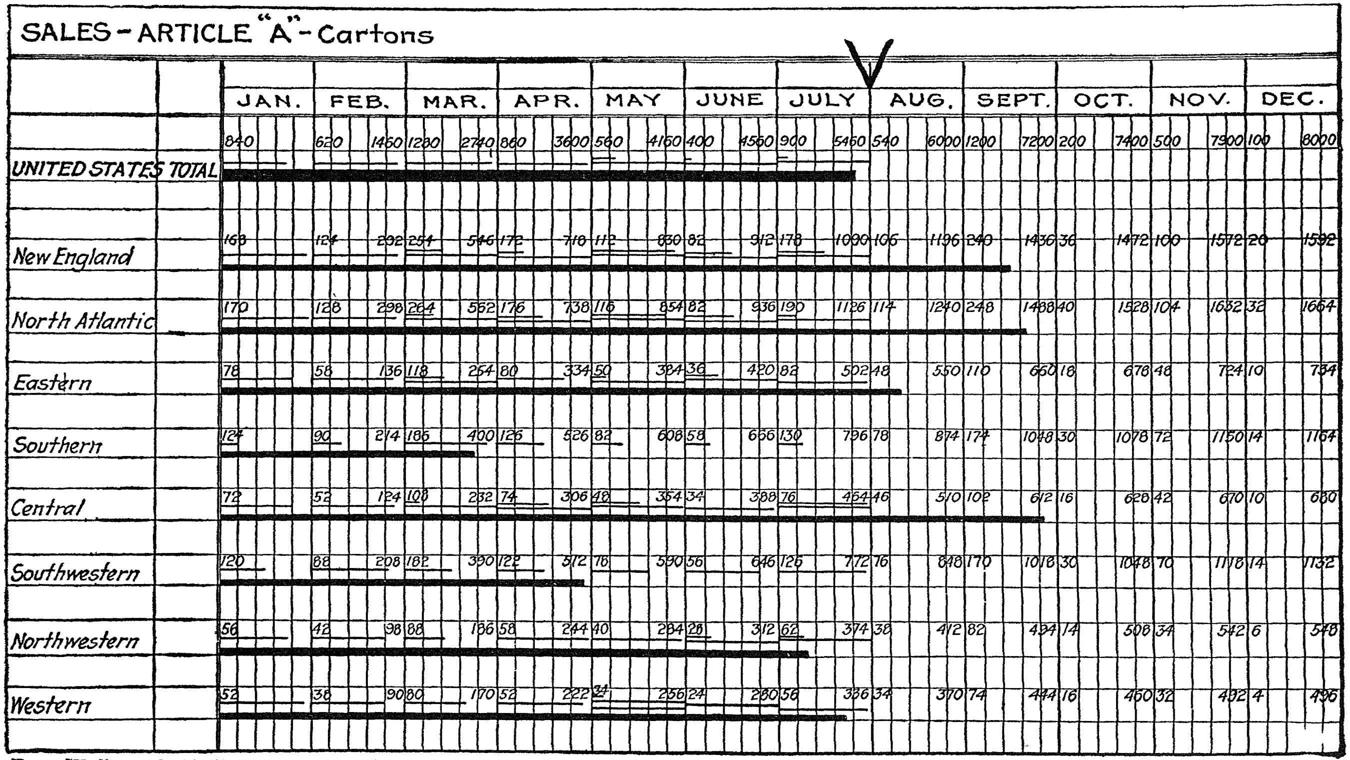


Fig. 238. The Chart on March 31st. The light lines show 150% of March quota done in March; the heavy line shows total done to date is one-sixth of a month ahead of schedule. The V-shaped mark (in pencil) shows date of last entry on chart.

true that the chart shows when you fell down, and you can, by notes, comments, or various symbols enter your excuses upon the chart, but first and last you cannot escape the fact that you fell down on the job. Likewise, if you have done better than your task (standard, quota, expectation, or what



From Wallace Clark's "The Gantt Chart", published by the Ronald Press.

Fig. 239. A Progress-Chart of Sales by Districts.

you will) the chart will bear emphatic witness to that fact; it will proclaim your success from the housetops.[2]

Picture to yourself a busy sales or factory manager receiving the usual detailed report on the production or activities of his various departments. The record is in tabular form showing what each department has performed during the month, or during the year to date. Before he can decide whether the work is satisfactory, he must study each figure, and in the light of his knowledge of all the various factors and circumstances, decide whether each item shown is satisfactory or not. It is a task which will cost him many hours of close concentration on every occasion when the report is submitted to him, and each time he will find difficulty in remembering just what were his previous decisions about each item. An executive's time is being taken up, not in getting things done, but in thinking about them. It would not be so bad if it could be accomplished once for all; the pity of it is that it has to be repeated every time a report is submitted to him.

Now let us help this executive by giving him a Gantt progress-chart at the beginning of the year or period for which the chart is to run. Let us sit down with him and ask him, once for all, to consider the various factors which in his judgment will be the basis of satisfactory work during the year to come, and in the light of those factors to determine upon a reasonable standard for the coming year's activity. Often he will give us merely the salient factors and their approximate influence and leave to us the working out of a detailed schedule in accordance with them. Often we will get much of the detail from subordinates closely in touch with them. In any case, what we are going to try to do is to devise for the entire coming period a schedule of reasonable expectation of the business for the coming period (e.g. year) worked out in detail for each element, department, or other subdivision, and for each unit of time (month, week, or day). This reasonable standard or schedule of expectation is sometimes called a "quota" or "task." It will often work very well as a quota, or basis of rating by merits or demerits, subject, of course, to unforeseen

[2] "Unlike statistical diagrams, curve records, and similar *static* forms of presenting facts of the past (Gantt) charts . . . are *kinetic*, moving, and project through time the integral elements of service rendered in the past toward the goal in the future."—Walter N. Polakov, *Principles of Industrial Philosophy*, Proceedings of the American Society of Mechanical Engineers, December, 1920.

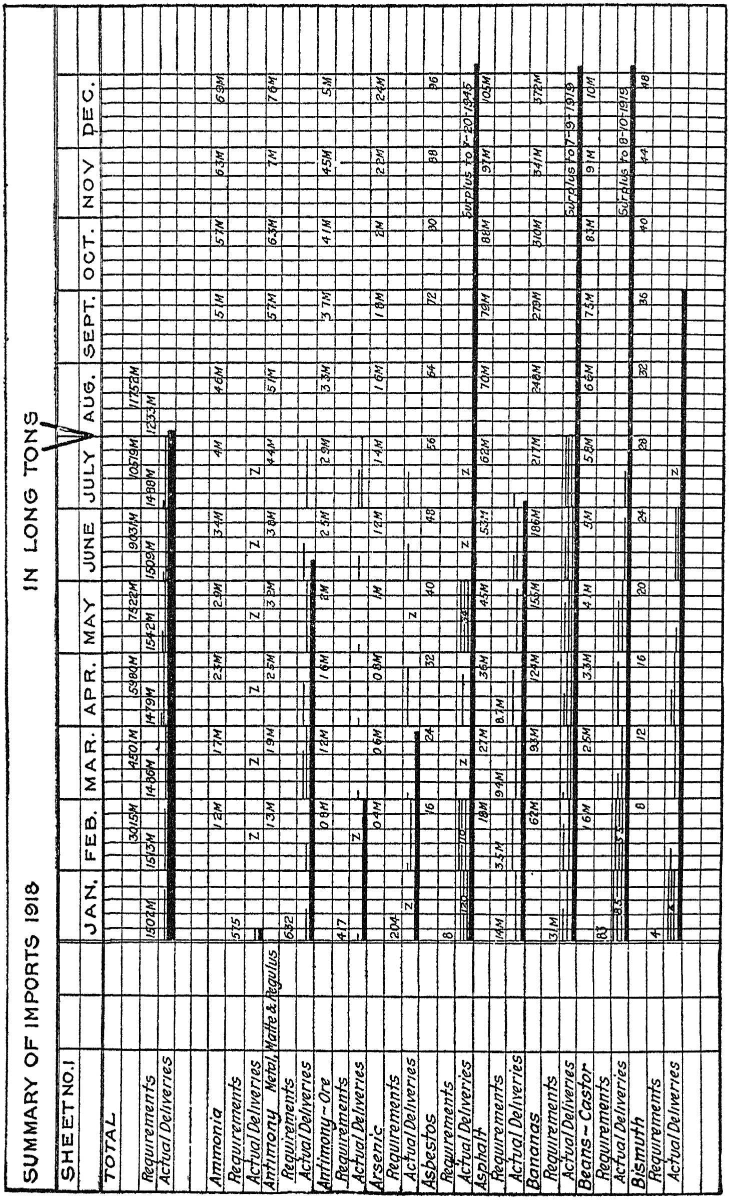


Fig. 240. A Progress-Chart Used For a National Inventory.
One of the charts kept by the Shipping Board during the war. This chart shows that the country's requirements of bananas for a year in advance had been imported, and that enough asbestos was brought in for the next twenty-seven years.—*From Wallace Clark's "The Gantt Chart", published by the Ronald Press.*

changes in the various factors or circumstances. But the name is not important, and whether we call this a quota, a standard, or a schedule, the point is that we have reached what the

executive considers to be the best basis for the judgment of work done. And we have finished the major part of the work of preparing a Gantt progress-chart. We enter these figures in the blank form and wait for performance to show the accomplishments as they occur, graphically on the chart.

Sometimes the difficulty of preparing a reasonable quota is so great, or the factors and circumstances which will be involved in the work are still so obscure, that a quota is not desirable. Nevertheless, you will find the executive later on using some figures or other for comparison. Most frequently he will be using the figures for the previous year as a basis of comparison. So when we cannot make up an ideal standard, we merely use the last year's record, or perhaps an average of several recent years, in the place of a quota on the chart. Gantt engineers, working on production and office problems, have developed a scientific technique in quota or schedule-making which is intimately connected with and greatly simplifies the problem of cost accounting in the plant. But whether these scientifically reliable quotas, or merely rough guesses, or simply the previous records, be used as the standard, it is all one to the progress-chart. The chart will use any basis adopted, and judge the performance in terms thereof.

Note the time-and-labor-saving value of the progress-chart. After the fundamental decision as to schedule standards have been made, the chart works automatically. It is a machine, passing up to the executive his own judgments. All the labor involved has been transferred to clerks. The executive merely glances down the chart, noting the length of the heavy bars. His attention is immediately drawn to the exceptional performances which he would wish to study. There is no dodging or forgetting these exceptional cases as shown on the chart, and the executive is enabled either to discount the exceptional cases in the light of further developments or unforeseen circumstances, or to take immediate action where such action is called for.

This method of charting is so simple that Gantt engineers are accustomed to install it in shops for the use of foremen, as well as for the central planning and executive departments. They are accustomed to enter the quota in ink and the graphic bars in pencil (black lead pencil). The maintenance of these charts requires no special staff of draftsmen or computers; they are used and filled in by the workmen and foremen them-

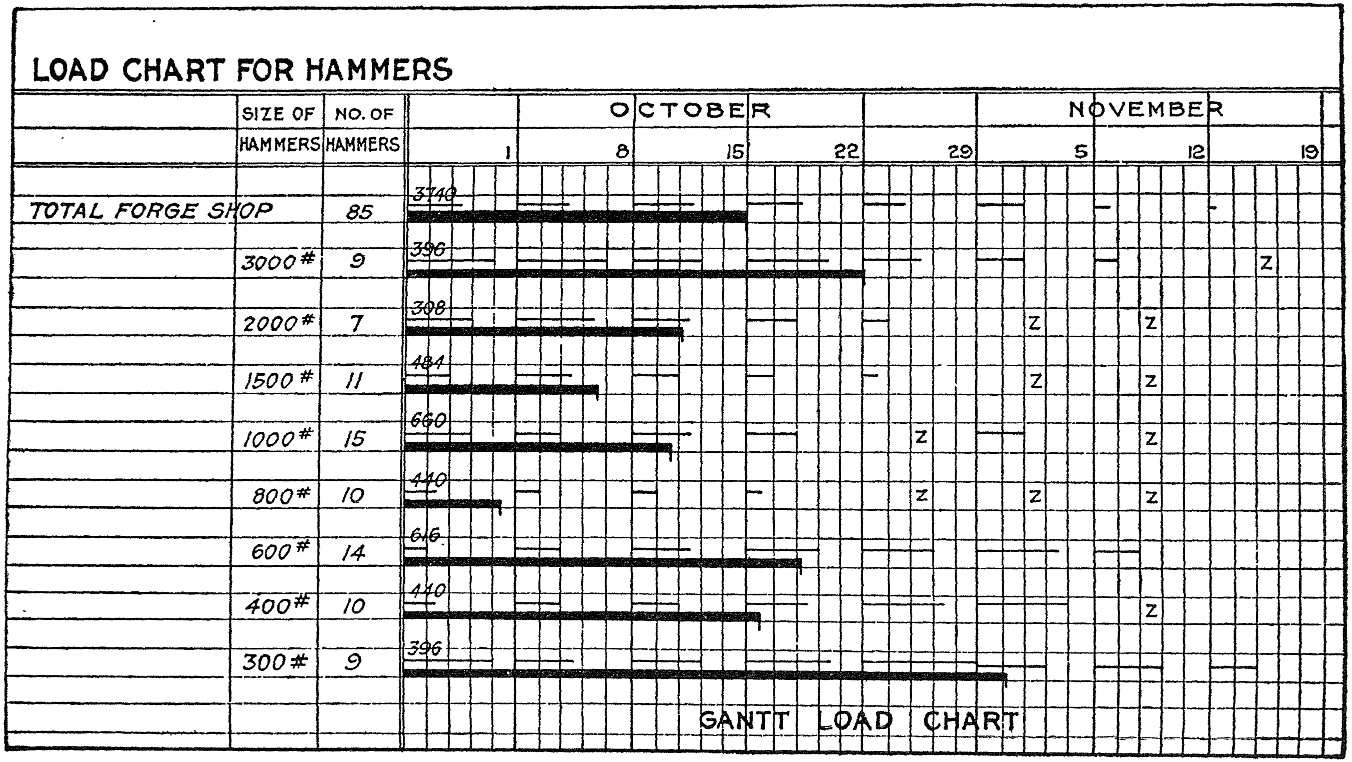
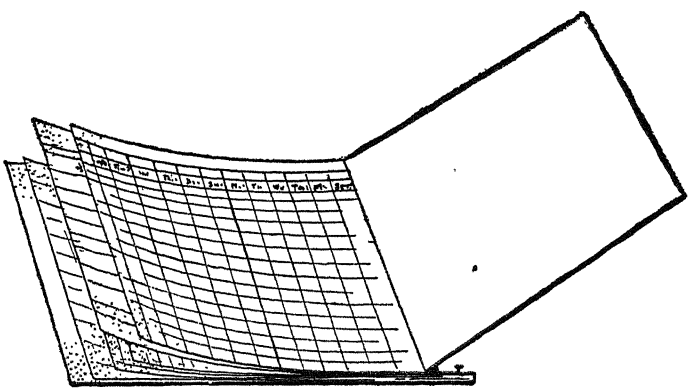


Fig. 241. Pencils Are Used in the Work-shop.

Detail of a quarterly form, with thirteen weekly periods and percentage subdivisions. Other forms are quarterly by weeks with daily subdivisions, and fortnightly by days with hourly subdivisions (the Saturday periods being narrower for the half-days).—*From Wallace Clark's "The Gantt Chart", published by the Ronald Press.*

selves, the bars being roughly sketched in without regard to nice appearance. Where the charts are used in a statistical department, or for the higher officials of a concern, it is perhaps better to observe more care in the appearance of the chart, and in this case colors can be used to advantage. The monthly or individual time unit quotas can be typed on the chart in black typewriting and the cumulatives in red; the light bars can be ruled in black India ink with a drawing pen, the heavy cumulative, in red waterproof ink. If this color distinction is observed, the ends of the red bars each month should be marked with light black cross-lines vertically cutting the red bar into segments, and showing the position of the cumulative at various times in the past. The initial letter in the name of the month can also be entered in black on the red bar to identify these points. When no color distinction is made and the cumulative bar is black, it is better to draw slight notches below each final position of the cumulative bar to keep this record of cumulatives. The first line upon a progress-chart is usually used for the total of the other lines and its bars are drawn with extra wide or heavy rulings, for the sake of emphasis.

A further refinement has been adopted by Gantt engineers in the binding of these charts. The charts are bound at the righthand edge of the paper in order that the stubs at the left-hand side may be quickly seen. This is of benefit where a large number of progress-charts are kept in bound volumes.



From Wallace Clark's "The Gantt Chart", published by the Ronald Press.

Fig. 242. Data on a Short Fly-sheet—the Gantt Way.

It has been said that the data should always be available with every chart and the Gantt progress-chart is no exception. Performance or accomplishment is portrayed graphically by the bars, but the figures which these bars represent are not shown on the chart. A blank form, similar in its ruling to the form of the chart, is laid immediately above the chart page with the column for stubs removed, so that the page will be short and the stubs as entered on the chart will be visible for both chart and data sheet. On this blank form the data of performance is recorded in writing, the figures for both individual time units and cumulatives being shown in the same spaces as used for their corresponding quotas on the chart. Because this method presents the data rather than the chart itself when the page is first opened, and the data sheet has to be turned over to read the chart-sheet, the writer's individual practice is to reverse the arrangement and place the chart on the short sheet and the data on the full sheet underneath—an act of heresy which Gantt engineers are not expected to endorse.

It is with reluctance that we leave the subject of progress charts. So far as the executive interested in the graphic control of operations is concerned, the entire subject of charts and graphs begins and ends with this chapter. No other method of charting has yet been invented which presents the essentials of operating so forcibly, so clearly or in such small space. The details for each department can be shown on departmental sheets with the total for the department at the top of each chart. The various department totals can be brought together upon a single plant sheet which will show to the plant manager at a glance his various departments, and the total for the plant, carried at the top of his chart. Likewise, on a single summary chart for the entire business, the president or controller can see the work of the various plants, with a total for the entire business at the top of his chart. In other words, this method of charting can be carried to the last degree of detail or reduced to the shortest possible summary, and the entire structure of an industry can be shown by a similar structure of co-ordinated charts. Once started, the work can be carried out almost entirely by clerks and secretaries, freeing the executives from routine, analytical work, and largely freeing any special statistical department for research work. The progress-chart is the "looking-forward"

chart par excellence, beside which all other charts are either research methods or "looking-backward" records.[3]

---

[3] The best descriptions of the Gantt charts are to be found in Mr. Wallace Clark's *The Gantt Chart*, Ronald Press, New York City.

The following articles in periodicals may also be consulted:

Polakov, Walter N., *Kinetic Statistics as an Aid to Production and Distribution*, Journal of American Statistical Association, Sept., 1922, p. 359.

Clark, Wallace, *Installing Gantt Production Methods*, Industrial Management, June, 1920.

The books of Mr. H. L. Gantt, *Organizing for Work*, *Industrial Leadership*, and *Work, Wages and Profit*, also contain mention of the charts.

CHAPTER XXIV

SUMMARY CHARTS

Historical data often comes in three sets of figures showing income (credits), outgo (debits), and balance. A wide variety of names is used for these three sets. Sometimes they are

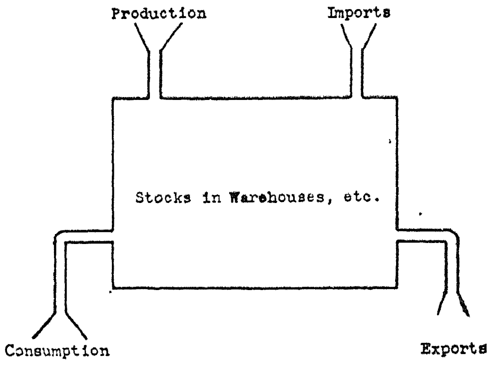


Fig. 243. The Flow of Goods.

called production, consumption, and stocks. Sometimes they are known as export, import, and foreign trade balance. Sometimes there are several groups of these three linked together in a single chain of events. Thus in a single business concern, the purchasing agent will keep a record of his orders (credits), receipts (debits), and balance on order. He will also probably keep a record of his receipts (credits), uses (debits), and balance on hand of raw materials. The factory manager may keep a record of withdrawals from raw material stock (uses) (credits), production (debits), and goods in process (balance). He will certainly keep a record of production (credits), shipments (debits) and finished stock on hand. If goods are stocked in warehouses, the warehouse clerk will keep a record of receipts (shipments in) (credits), sales or consignments (shipments out) (debits), and stock on hand.

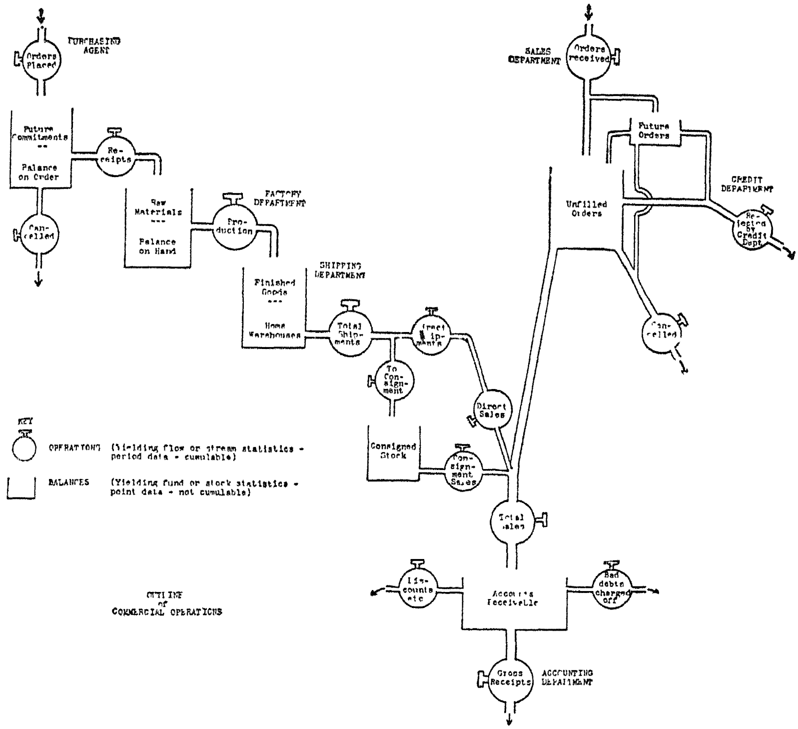


Fig. 244. A Typical Industrial Process.

The accounting department will keep track of sales (credits), payments received (debits) and accounts payable (balance due). Other steps may be inserted between the above. In these chain periods of groups of figures, the outgo (debit) from one deposit station will appear as the income (credit) to another.

In fact all historical data belongs to one or another of these three classes. Population statistics are merely a balance between the number of births (income) and the number of deaths (outgo) in a community or country. Crop and production statistics generally belong to the income class, the figures for consumption and goods in storage or process being rarely available. Examples might be multiplied.

The economists' distinction between stocks or funds of goods and streams or flows of goods goes to the bottom of all this. For if you will regard the stocks of goods as a reservoir or body in repose, you will see that the income is a flow or stream of goods into this reservoir and the outgo is a flow or

stream of goods out of this reservoir. And if you are mathematically inclined, you will notice that whenever any two of these three sets of data are given us, we can easily compute the third. If we know the January 1st inventory, the production during the year, and December 30th inventory, we can easily compute the shipments or sales for the year, that is, the withdrawal or outgo.

Now there is a very important distinction between stocks and streams (whether income or outgo streams). The former, stocks, can only be measured at a point of time, while the latter, streams, can only be measured during a period of time, between two points of time. In the language of physicists, streams of goods have one more dimension than stocks, for they have the added dimension of time. And the result of all this is that while you can cumulate or total up the stream figures, you cannot cumulate or total up the stock figures. You cannot cumulate daily the population of a city in order to get the population monthly, for population is a stock figure, though you could have cumulated the number of births daily to get the number of births monthly. You cannot cumulate your monthly balance in order to get at your annual balance, but you must cumulate your monthly production in order to get at your annual production. The distinction between stock figures and income or outgo figures is fundamental.

The usual method of showing these three sets of figures is to use three curves upon a single chart. The use of three curves for such dissimilar data is always confusing to the reader of the chart, as the thin plotted line of the curve represents in one case a static or stationary stock of goods and in another a series of separate and distinct additions or subtractions. If the chart shows monthly data, a recasting of the chart on an annual basis would raise the production and consumption curves to twelve times as high a level on the chart, but would leave the stock or balance curve unaffected.

The author has designed what, for the lack of a better name, he is accustomed to call a summary chart to show these three sets of figures together upon a single chart.

This summary chart is a combination of two vertical-bar charts, and a curve chart in which the bars show the stream figures (income and outgo) and the curve shows the balance or stock figures. This distinction makes clear to the most

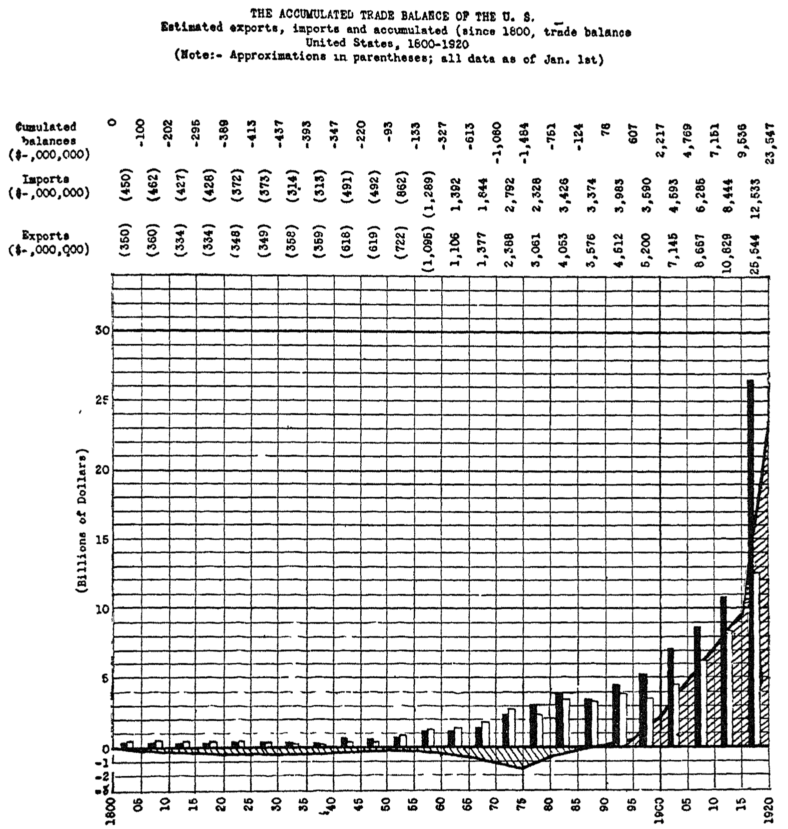


Fig. 245. The Summary-Chart.

casual reader that the total income is made up of all the little income bars and the total outgo is made up of all the little outgo bars while the stock of goods changes according to the fluctuations of the curve. In practice it has been found that the chart succeeds admirably in showing simply and clearly the rather complicated relations of the three sets of figures. And the chart is more or less unique in its nice use of both bars and curves simultaneously.

A color distinction is made between the income and outgo bars, income being black or green and outgo red, or income being white and outgo being black. A pale tint, gray or half-tone, is given to the entire area between the curve and the zero line, except where the bars cross this area. The stock curve or balance curve being plotted at the beginning and end

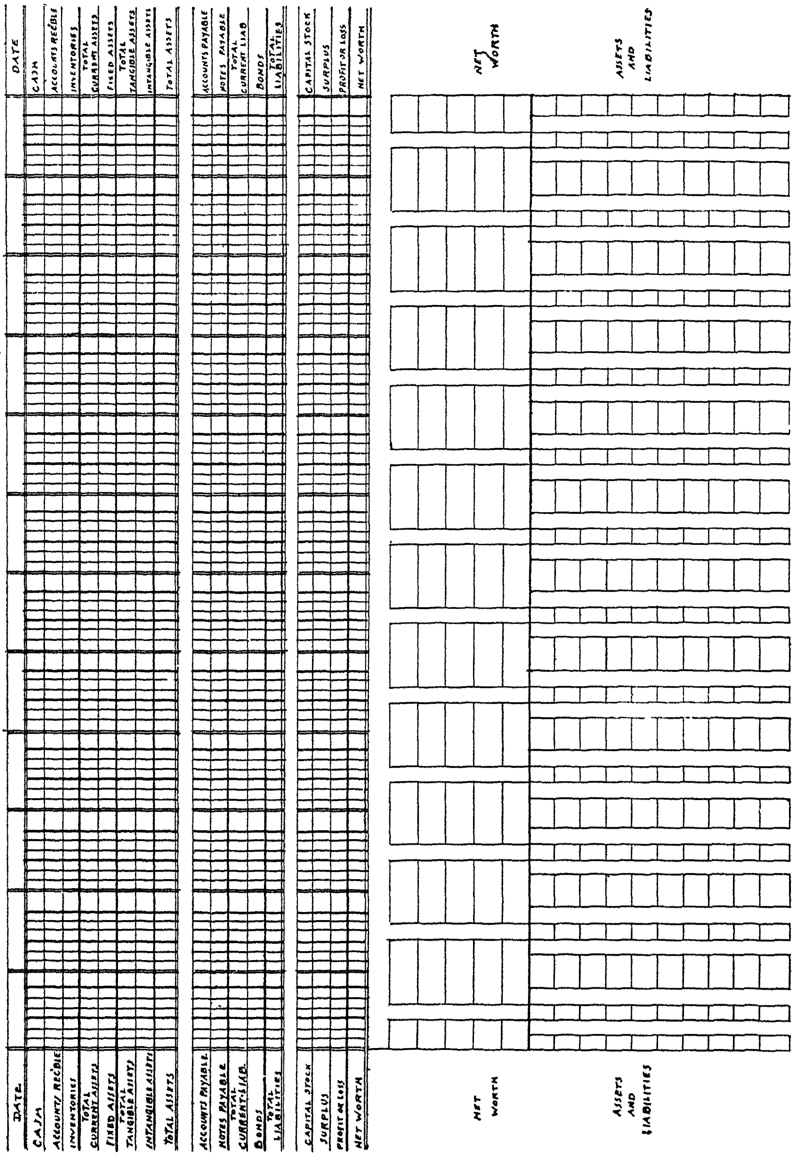


Fig. 246. A Net Worth Chart-Form.

This chart-form has open spaces for two bars, (for assets and liabilities) downward and one bar (for net worth) upward, in each month, together with actual record above. The bars can be segmented as indicated by the data. As this chart is more effective in colors, only the form is shown here. Here, bars are used in the place of curves.

of each period of time, since these are the dates of inventories, the plotting points for the curve are the ordinates of the various points of time. The vertical bars on the other hand are placed within the spaces for the periods of time (days, weeks, months or years) between the ordinates. Two bars appear in each space, the first for income and the second for outgo. These bars should not be any wider than is necessary to make them quite clear. The remaining space between bars is used for the plotting of the curve and for the shading of the area under the curve. In its finished form the curve, or stock figure, of balance on hand, which might be called a band curve because of its shaded area, forms the background of the chart. Against this background, the quantity added to the stock each period of time, and the quantity subtracted therefrom, that is, the bars or stream figures, of income and outgo appear in the foreground as solid bars.

The vertical scale for the summary chart (like the horizontal scale) should be the same for all three sets of figures. When this uniform scale is used it will be easily seen that the changes or fluctuations of the stock curve exactly coincide in vertical distance with the difference between the lengths of the two stream bars. If the income bar is higher than the outgo bar, the curve of stock on hand will rise by the difference in height or surplus, and vice versa, if the outgo bar is the larger, the curve will fall by the difference or deficit. Another advantage of the uniform scale is that the height of the stock curve can be compared with the height of the various outgo bars to show easily the approximate period of time the stock on hand would suffice if income were to cease. Where the stock represents invested capital, a firm desires to keep this margin of stock on hand as low as possible, and the chart shows clearly the size of the inventory compared with the periodic requirements. In the case of non-perishable goods or semi-perishable goods, such as fresh food stuffs, the storage, or amounts withdrawn from circulation, vary a great deal seasonally, and warehousing conditions must be sufficient to meet the maximum stock which will be in storage.

The summary chart can be made to carry a great deal of detail by converting the bars into compound or segmented bars, and the curve into a band-chart or segmented curve. In this case, care must be used in the shading of the segments of the bars, in order that they will not obscure the primary

distinction between income and outgo bars. It is, however, possible to make all the shadings of the black or green income bar in various degrees of intensity by various black or green cross-hatched patterns, and to accomplish the same for the segments or layers of the red outgo bar and the green or black balance curve. The segments would indicate the various parts of the income, outgo or balance. When much detail is shown in this way, it becomes necessary to omit a portion of the data

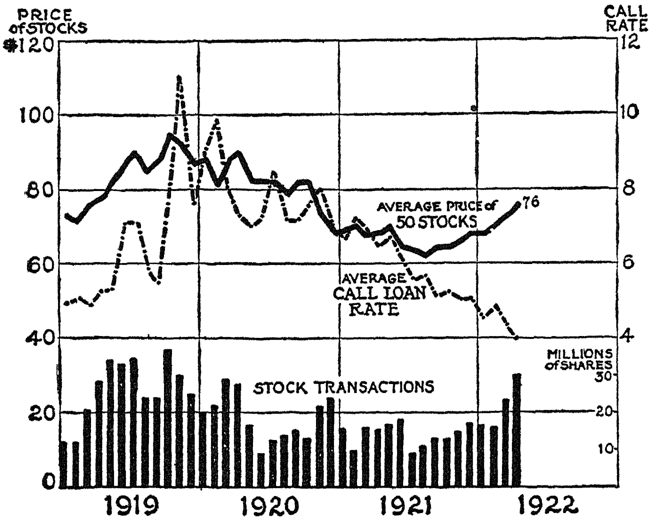


Fig. 247. A Customary and Sound Combination of Bars and Curves. Average price of stocks, average call loan rate, and total transactions in stocks each month.—*Permission of Mr. Carl Snyder.*

from the chart, or to adopt such wide intervals between ordinates that the data can be entered horizontally. The latter method usually calls for a chart of more than usual size.

The successful use in this chart of a distinction between vertical bars and curves suggests that a general rule could be made for all cases of complicated charts, wherever more simplicity is desired. This rule would be that bars should be used for streams of goods, and curves should be used for stock figures. The rule would have many exceptions owing to the convenience of the curve form in general, but it would appear to be a sound principle to follow, wherever the choice is open to us and the use of either curves or bars alone is not felt to give sufficient clearness.

CHAPTER XXV

SILHOUETTE BAR-CHARTS

For certain purposes, only a few details of historical data are required for each one of a large number of different and heterogeneous phenomena. In price movements or stock quotations, for example, the practical man is interested

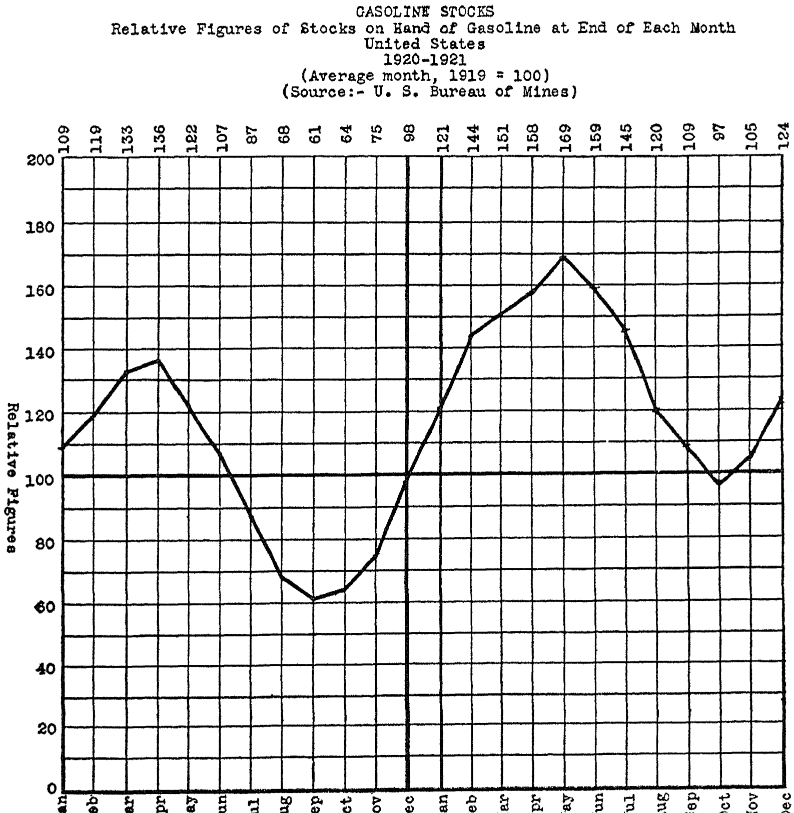


Fig. 248. A Curve is too Detailed and Large.

	RELATIVE PRODUCTION (1919=100).							
	Maximum since end of 1919.	Minimum since end of 1919.	1920 average.	1921 average.	Feb., 1921.	Mar., 1921.	Feb., 1922.	Mar., 1922.
FOODSTUFFS:								
Wheat flour	125	64	82	91	64	82	88	89
Beef products	109	67	92	83	67	83	75	......
Pork products	151	58	93	97	114	92	102	......
Lamb and mutton	110	58	80	94	89	102	70	......
Sugar (meltings)	165	40	104	92	80	133	128	165
Oleomargarine [2]	126	26	103	60	70	73	42	52
Cottonseed oil	349	7	100	166	247	229	140	110
Condensed milk	121	20	76	71	34	58	......	......
Butter	177	64	99	118	76	91	......	......
Cheese	169	41	86	83	49	68	......	......
Ice cream	468	42	111	153	44	71	......	......
CLOTHING:								
Cotton (consumption)	114	57	109	79	76	84	91	100
Wool (consumption)	126	42	83	95	64	83	111	124
Sole leather	95	63	82	79	63	72	78	78
FUELS:								
Anthracite coal	119	63	101	99	105	101	92	119
Bituminous coal	137	74	121	89	81	79	107	131
Beehive coke	127	11	110	30	55	36	35	46
By-product coke	......	62	122	79	90	85	86	102
Crude petroleum	149	104	117	124	112	130	130	149
Gasoline	141	98	123	130	118	127	121	......
Kerosene	110	71	99	83	84	87	86	......
Gas and fuel oil	136	93	116	127	115	119	120	......
Lubricating oil	135	89	124	104	103	103	98	......
Electric power	119	98	113	105	98	105	107	117
METALS:								
Pig iron	132	34	119	54	76	63	64	80
Steel ingots	140	34	121	59	74	66	74	100
Copper	83	17	94	37	71	83	35	58
Zinc	126	38	105	47	46	41	59	69
Silver	129	80	100	95	116	129	82	89
Gold	181	79	88	113	93	100	94	99
TOBACCO:								
Cigars [2]	128	75	112	96	84	95	76	90
Cigarettes [2]	116	64	84	96	93	101	71	92
Manufactured tobacco [2]	119	50	94	91	85	100	92	108
LUMBER.								
Yellow pine	113	69	94	99	88	101	98	113
Western pine	121	20	121	67	20	57	38	53
North Carolina pine	153	33	98	88	63	71	149	153
California white and sugar pine	204	8	121	78	11	12	19	15
California redwood	156	57	122	109	92	120	90	135
Douglas fir	118	44	102	79	57	68	108	107
Michigan hardwood	111	32	86	60	68	86	49	49
Northern hardwoods	161	21	105	88	117	147	72	118
Hemlock	120	33	91	57	57	52	44	67
Oak flooring	202	42	106	123	55	84	171	202
PAPER:								
Newsprint	114	69	110	89	90	94	85	103
All other paper	132	69	121	86	76	83	101	119
Mechanical wood pulp	143	55	109	87	98	118	82	119
Chemical wood pulp	138	64	117	79	78	74	90	106
Corrugated paper board [3]	129	30	104	65	42	49	86	100
Solid fiber paper board [3]	142	18	104	89	53	75	100	116
STONE, CLAY, AND SAND PRODUCTS:								
Silica brick	130	13	106	40	66	63	47	65
Clay fire brick	127	43	120	63	81	83	68	84
Face brick	121	34	100	100	34	41	51	93
Cement	157	61	125	122	65	101	64	100
Glass bottles	124	48	104	69	87	68	81	......
BUILDING EQUIPMENT:								
Baths, enamel	189	65	149	120	71	78	152	189
Lavatories, enamel	199	86	112	127	136	129	154	199
Sinks, enamel	170	80	110	122	96	128	135	166
Buildings (contracted for)	118	30	72	70	36	58	65	112
TRANSPORTATION VEHICLES:								
Automobiles, passenger	[1] 121	[1] 51	114	93	......	......	79	111
Motor trucks	[1] 52	[1] 32	102	46	......	......	49	74
Locomotives	135	13	89	50	79	72	20	17
Ships	79	([2])	67	30	32	42	11	2

[1] Since July 1, 1921.
[2] As represented by tax-paid withdrawals.
[3] Relative to last 6 months of 1919.

From Monthly Survey of Current Business.

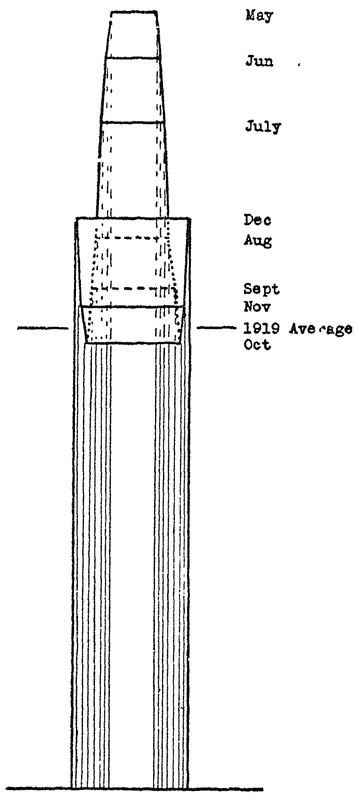
Fig. 249. The Essential Data.

primarily in the latest quotation, but would also like to know whether this last quotation is an increase or a decline from the quotations on previous dates, and how it compares with past maximum and minimum prices. Here, then, are four points of interest to him, the present, the immediate past, and the prior record-making peak and valley prices (regardless of the dates or time of the latter). And the point is that we want to see these facts, not for one only but for a large number of commodities. It would be easy enough to present curves for the individual commodities, or even to combine a few on one curve-chart, but how can we present only the facts wanted, for all the commodities, graphically in one simple chart?

If you were to stand a historical curve-chart up on edge and view it from the side toward which the curve is moving, you might succeed in imagining that the curve was really snaking its way directly at you. And if it had actual volume, instead of being a thin line of ink, you would see most clearly its nearest end representing the last value, and behind that a short portion of its previous values, and still further back you could make out the silhouette of its extreme peak and valley points. Eureka! These things are all you wanted to know about each individual curve, and seen from the planes in which the curves lie, each curve compresses to the width of its imaginary columns or vertical bars. This suggests the method of graphing to which, for lack of a better one, we give the name of silhouetting. The silhouette curve-bar is a recent development in graphics and probably has not yet reached its final stage.

Since the graph is really a projection of a large number of curves shooting straight out of the page toward the reader, something must obviously be done to lift the nearest ends of the curves from their other and earlier positions. The method of segmented bars alone gives too much flatness to a picture which is really a projection of three dimensions on two—a condensation of a three dimension model into a two dimension sheet. For this reason it is obvious that the nearest ends of the chart stand out clearly and appear to be wider, precisely as if photographed from a real model. The effort here being to produce the effects of perspective, the portion of each bar which represents its latest reading or value should be of full width, but the earlier readings, and in particular the past peaks and valleys, should be considerably narrower, to give the effect of greater distance.

If the portion of each bar connecting the latest and the next previous values be kept of uniform width, it is necessary to show whether the change has been one of rise or fall, that



GASOLINE STOCKS
Relative Figures of Stocks on Hand of Gasoline
at End of Each Month
United States
1920-1921
(Average month, 1919 = 100)
(Source:- U. S. Bureau of Mines)

Fig. 250. The Same Curve Seen From Its End.

is, which end of the bar is the latest reading. This can be indicated with a small arrow-head in the bar, or by solid shading of one color for rises and of a totally different color for declines, or by both methods together. Moreover the latest reading might be indicated by a star or other symbol which the reader can quickly glimpse. On the other hand, if

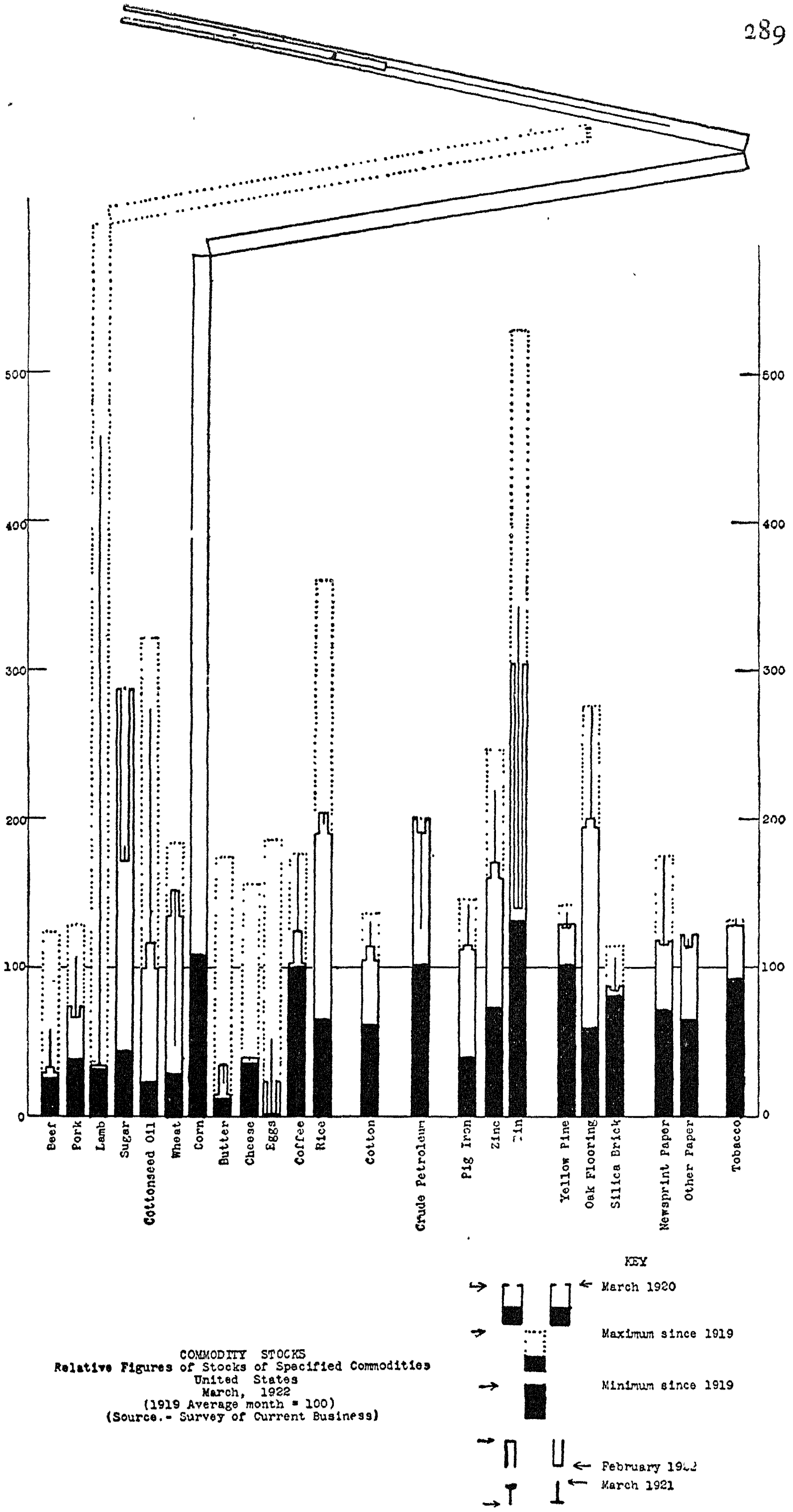
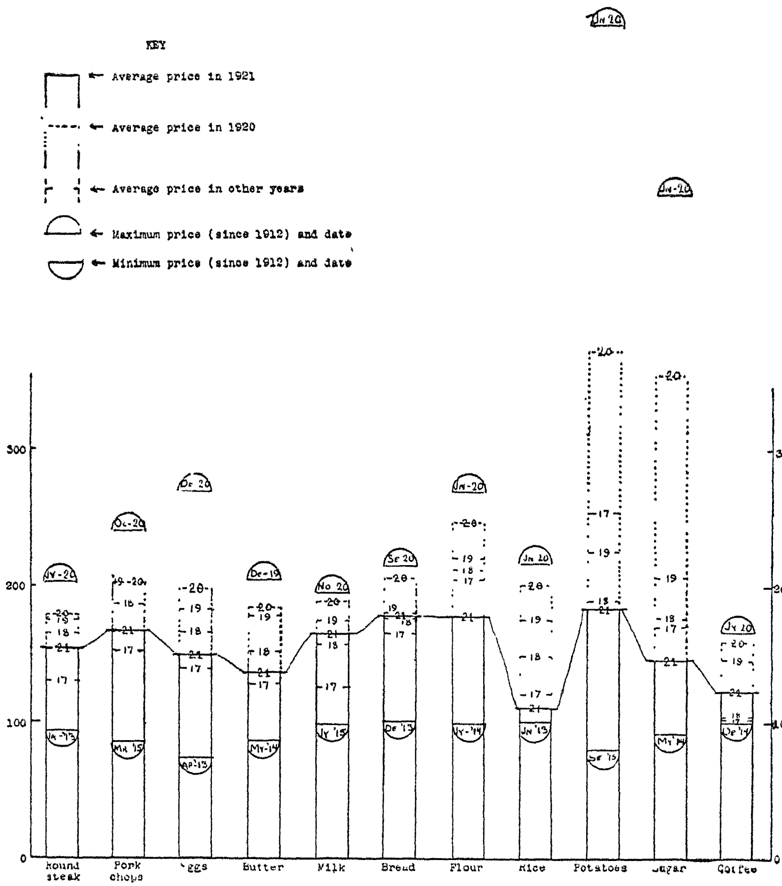


Fig. 251. A Detailed Form.

we are willing to make the bar of tapering shape, then the narrower end could indicate the earlier, and the wider end, the later value.

A more pictorial method may, however, in some cases prove to be the more efficient means of flashing the story of the chart to the reader. Thus, if a large circle be used for the latest value, and a small circle for the next previous, with a triangle or star for the highest past peak and an inverted



RETAIL PRICES
Relative Figures of Annual Average Prices at Retail of Specified Commodities
United States
1917-1921
(1913 average = 100)
(Source - United States Bureau of Labor Statistics)

Fig. 252. Data in the Chart.

triangle or square for the lowest past valley, it would seem that the results would be easily understood at a glance. Each of the four values could be strung upon the same central line (or ordinate) serving as a connecting thread in the place of the bar. The two circles for recent values could be white for rising values and black for declining ones. Such a pictorial system would make possible the addition of even further symbols, such as still smaller circles for the second previous reading when this was sufficiently different from the last two readings, and smaller triangles for minor peaks and valleys. The peak and valley symbols could contain numerals representing their years or approximate dates. The enterprising reader will find still other embellishments, which, so long as they increase the "visibility" of the facts or the speed with which they are flashed to the accustomed and unaccustomed eyes, will be justified.

The labelling of the various bar-curves in this chart calls for great care, particularly when the data is heterogeneous. Each compressed curve should be so distinctly and clearly

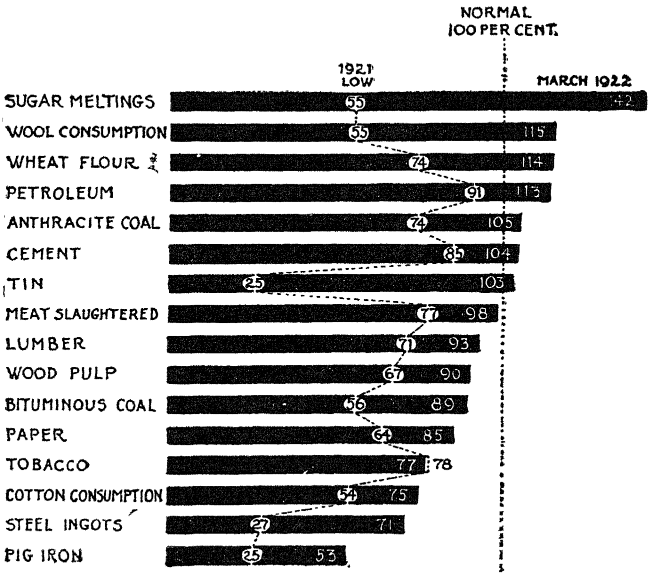
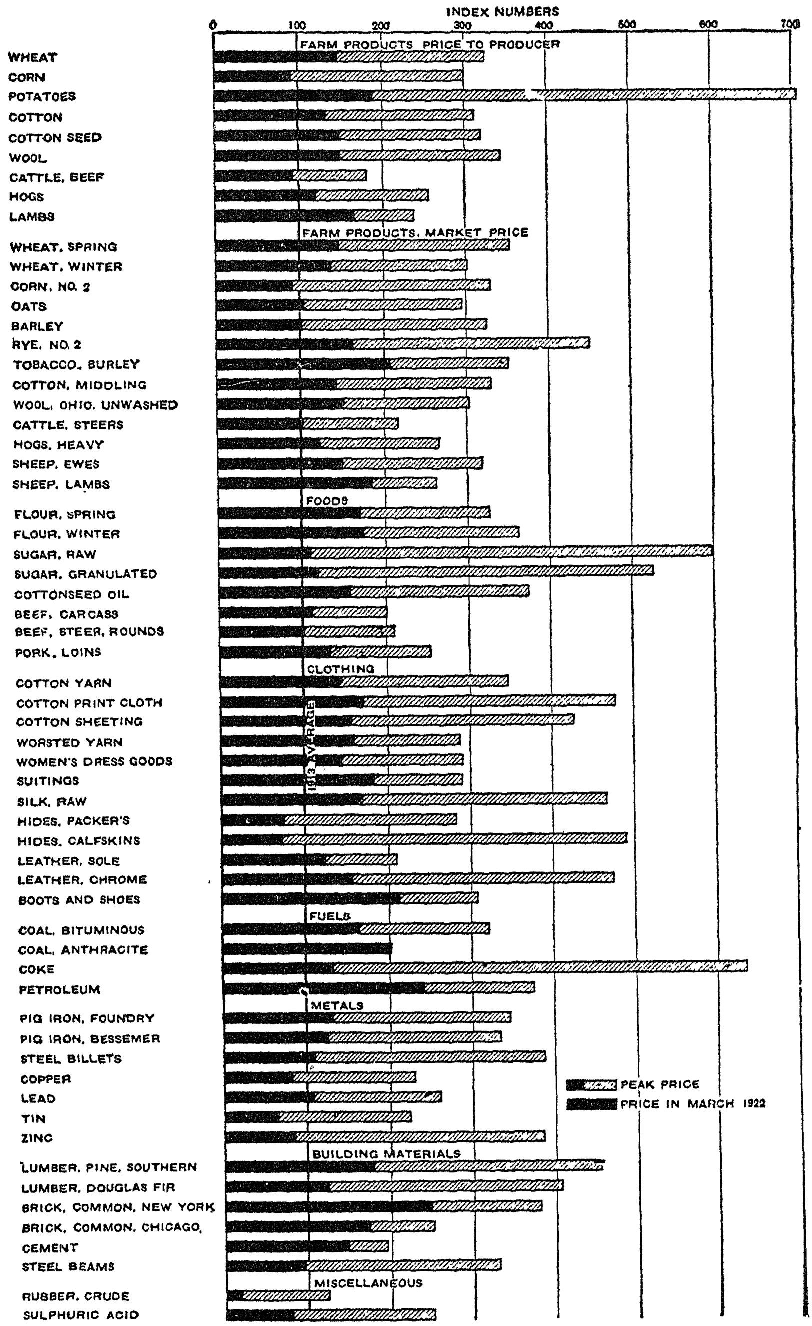


Fig. 253. Simple Silhouette Bars Presented Horizontally.

Production of basic commodities in March 1922, and the low point in 1921 compared with normal production. In cases in which March production figures are not available, February figures are shown.—*Permission of Mr. Carl Snyder.*

COMPARISON OF PRESENT WHOLESALE PRICES WITH 1920 AND PRE-WAR.

(Relative prices 1913=100.)



From the Survey of Current Business.

Fig. 254. A Silhouette-Bar Chart Set Horizontally.

labelled by its descriptive title at the base of the chart immediately underneath it, that the reader may have no difficulty in finding the graph for any particular item in the series, or in finding the title of any particular graph. Because these charts generally contain a very long list of items, it is well to divide them into smaller groups with slight margins between the groups, and it goes without saying that these groups should be as logical and useful to the reader as possible.

The silhouette curve-bar can be projected on either an amount-of-change or a rate-of-change scale. Usually the index numbers are used in the place of the absolute numbers, so as to make the items and graphs comparable, and usually the scales are arithmetical. But index numbers can equally well be shown on logarithmic scales and the latter give an added refinement to the chart which shows with greater accuracy the significance of changes, to those readers who understand it.

CHAPTER XXVI

INDEX NUMBERS

It is one of the most important functions of the statistician to compare the behavior of different phenomena and find out whether there appears to be a relation of cause and effect between them. For if such a relation exists, we ordinarily expect to find evidence of it in their fluctuation. If one phenomenon is directly or indirectly the cause of the other we may expect to find the fluctuations of the first paralleled or mirrored in the fluctuations of the second. If both are the effect of a third common cause, we may still expect to find a marked similarity between their movements. Often there is a delay or lag between the time of the movements of one object and the reaction upon the movements of the other. Thus it has been shown that the fluctuations in the production of pig-iron follow closely those in the corn crop after a lag of about two years.

The science of business forecasting is largely built up on these comparisons. Thus if pig-iron activity follow the corn crop exactly after a two years lag, it is easy to see that with a knowledge of corn prices today we would be able to forecast the prices of pig-iron two years from today. Some of the ablest statisticians in the country are engaged upon research in this forecasting problem. Unfortunately the relations are not simple and easily established, and the available information is not nearly complete enough at present to make general business forecasting very successful. It is, however, sometimes possible for a capable mathematician to construct a very accurate forecast for an individual business or industry in which the determining factors are more easily ascertained and measured.

When a condition of similar fluctuations (either parallel or mirrored) exists, the word "correlation" is used for the condition by both mathematicians and statisticians. They have a

complicated, mathematical process or formula for computing the degree of this correlation between two or more series of figures. The degree or extent of this correlation they indicate by a "correlation co-efficient." The mathematical work involved in determining this correlation co-efficient, which measures the degree of correlation between two series, is long and complicated. There is, however, a very simple method of detecting the existence of correlation or similarity of fluctuation between two series, which consists of plotting the curves of the two series and comparing the curves by sight. Do the wiggles in one curve parallel or mirror the wiggles in the other? If the curve has been properly plotted, correlation provided it exists, will be apparent at a glance. By using very translucent paper, you can subject the two curves to "light analysis," that is, you can lay one curve over the other, hold the two of them up to the light, and immediately see the slightest deviation of one from the other. The method does not give the correlation co-efficient, or exact measure of the degree of correlation, but it serves to give a sufficient idea of the amount of correlation for most purposes.

The only problem is to plot the two curves correctly. Ordinarily, the different series of data have different units of measurement. Thus corn is measured in bushels and pig-iron in tons, one a measure of volume and the other of weight. So far as prices are concerned, they read in common units of value, but one may lie much higher up on the scale than the other curve, because one may be measured in dollars and the other in cents. And as you know, the same fluctuation will be greatly magnified in a curve lying higher up on the chart, than in one positioned low. The fluctuations might be identical and yet when plotted on the same scale of numbers they would appear very dissimilar, because of the exaggeration of the fluctuations in the curve of iron, lying higher upon the chart.

There is however a very simple method of reducing two entirely different series to a common scale with a similar position and range on the chart. The trick is to use "index figures." In a previous chapter, the distinction between absolute and relative figures has been pointed out. The bushels of corn are absolute figures; but the percentages which these figures are of figures at a certain point of time are relative figures, also called index figures, index numbers and indices. When using

indices for a historical series, it is necessary always to state what year or time is considered as the basis for the percentages, that is, which year or time is taken as one hundred per cent. The hundred-per-cent year is called the base-year or "base" and the price or value at this time is called the "base figure" for the relative series or index figures.

The first step in reducing a series of data to index figures, therefore, is to select the base. Obviously a great deal depends upon the base you choose. If you select the highest

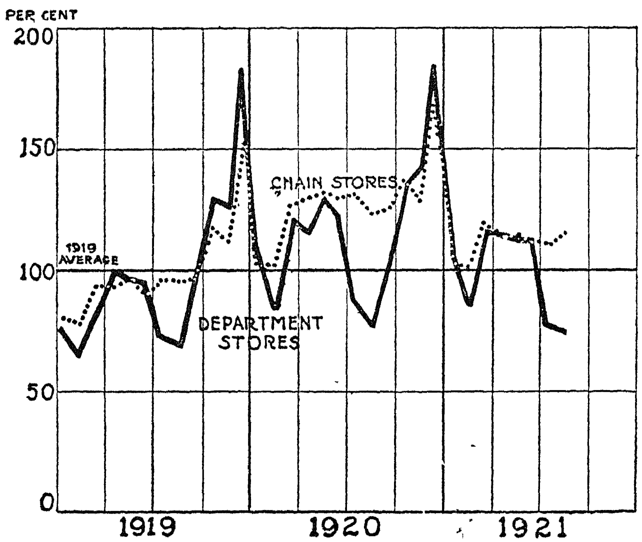


Fig. 255.

Sales of 57 department stores in the Second Federal Reserve District and 8 leading chain stores doing a country-wide business (average monthly sales in 1919 = 100%). *Permission of Mr. Carl Snyder.*

figure in the series the rest of the series will lie below the 100% line on the chart; if you select the lowest figure in the series, the entire curve will lie above the 100% line. The common practice is to use an average of a number of figures during times which were considered normal. Thus in its "price relatives," or index figure of prices, the Bureau of Labor Statistics has adopted the average price for the year 1913 as the base in all its price-series on the general theory that this was the last normal pre-war period of time. Others have adopted other periods of time as the bases for their series. In special cases you may have to adopt a special year regardless of its nor-

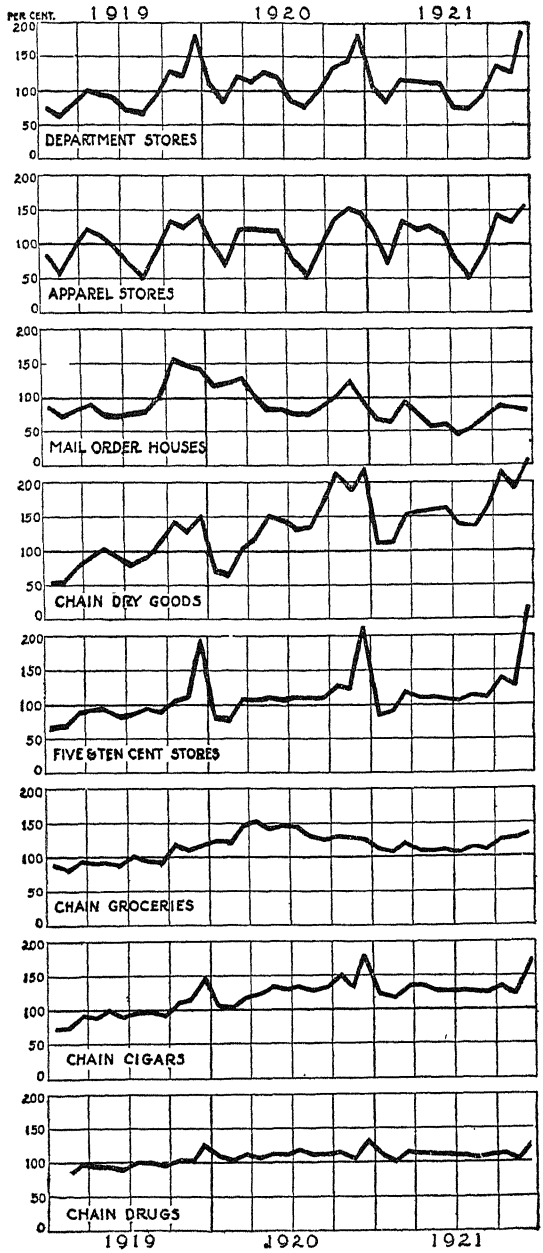


Fig. 256.

Monthly sales of department and apparel stores in the Second Federal Reserve District and of mail order houses and chain stores doing a country-wide business 1919 average = 100%).—*Permission of Mr. Carl Snyder.*

malcy. In comparing the index figures from different sources, you must convert both series to a common base year. A relative series can be changed from one base to another in the same way that it was constructed from the absolute series, that is, by dividing the series through by the value for the new base period.

PRODUCTION OF MANUFACTURED GOODS

Physical Volume of Production and Growth of Population
United States, 1899-1919. (Source:- from Mr. E. E. Day)

	1899	1900	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919
All manufactured goods	100	101	112	124	126	123	144	155	156	133	163	166	158	181	187	171	187	219	215	214	195
Textiles	100	101	102	116	116	116	129	130	132	121	149	141	130	148	156	157	168	208	201	203	191
Food	100	101	123	117	125	125	141	142	132	136	144	134	140	150	151	153	170	151	160	194	191
Population	100	102	104	106	108	110	112	115	117	119	121	123	125	127	130	132	133	135	137	138	140

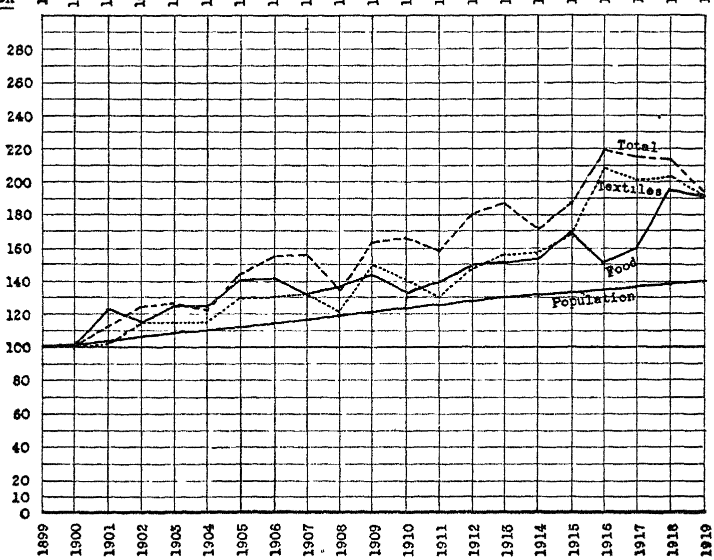


Fig. 257. Obviously, only Index Numbers are Possible.

Many books have been written on the subject of index numbers.[1] There is nothing difficult about the task of constructing relative figures for a single series of data. As already stated, you first select a base for the series and then divide

[1] The literature on this subject is considerable. In particular, the student should refer to the works of Wesley C. Mitchell; also to Irving Fisher, "Best Form of Index Numbers," American Statistical Association Quarterly, March, 1921, p. 533.

all the other figures of the series through by this base figure, turning them into percentages of the base, that is, into a relative series. To compare two series on a chart, you merely turn them both into relatives to the same base, then plot the curves of the two series and compare their fluctuations. The difficulty comes when you want to make a common index series for two relative series. Thus, if you have the price-series of various grades of steel, how will you make a single index series of figures for steel of all grades, that is, how will you combine these various relative series into one single index series. For if you are going to compare steel and wheat prices, it is obviously not an easy matter to have to compare a thousand different grades or kinds of steel with as many different grades of wheat. It is much easier if you have a single index figure for the price changes of all steel, and another for the price changes of all wheat. The problem of finding a series of figures which

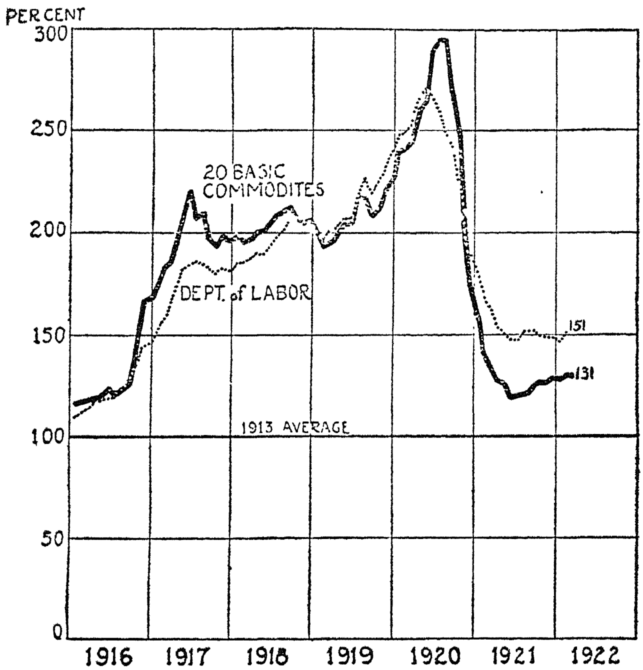


Fig. 258. Various Indices of the Same Phenomenon Produced by Different Methods of Weighting.

Index of the prices of 20 basic commodities compared with the Department of Labor Index (325 commodities).—*Permission of Mr. Carl Snyder.*

will serve as index numbers for several relative series is not an easy one.

Briefly, an index for a number of relative figures must be some sort of an average of those figures. But there are several kinds of averages, each with its own particular merits and purposes. For simplicity, let us suppose that we have only two original series, the price of a loaf of bread in the city of Oshkosh, and the price of a loaf of bread in Kalamazoo, and wish to find a single index for the price of a loaf of bread throughout the county containing these two towns (assuming that they together comprise the total population of one county). If

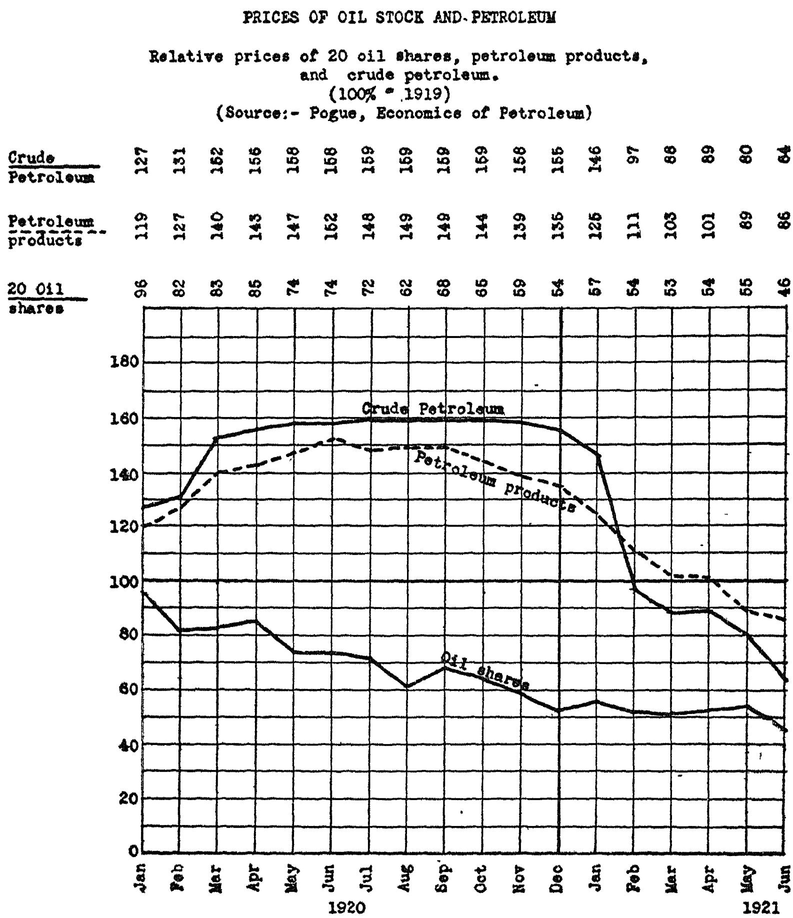


Fig. 259.

Kalamazoo bread sells at 10c a loaf and Oshkosh bread at 15c, the average price would appear to be 12½c. This is the simple arithmetic mean of 10 and 15c. If we are using a geometric mean as an average, the average would be around 12c, and if we are using a harmonic mean as the average, the average would be around 13c. For most purposes, however, the arithmetic mean, that is, the common or garden variety of average will do.

But let us suppose that Kalamazoo has a population of 900 persons and Oshkosh only 100 persons. Hence, for every loaf eaten in Oshkosh there will be nine loaves eaten in Kalamazoo and the true average price of every ten loaves will be about 10½c. A little study will show that the average loaf is nine times a Kalamazoo 10c loaf for every single time it is an Oshkosh 15c loaf. In other words, we must "weight" the figures before averaging them. Now this weighting is sometimes a very difficult problem. How would you combine changes in the cost of butter and changes in the cost of other foodstuffs

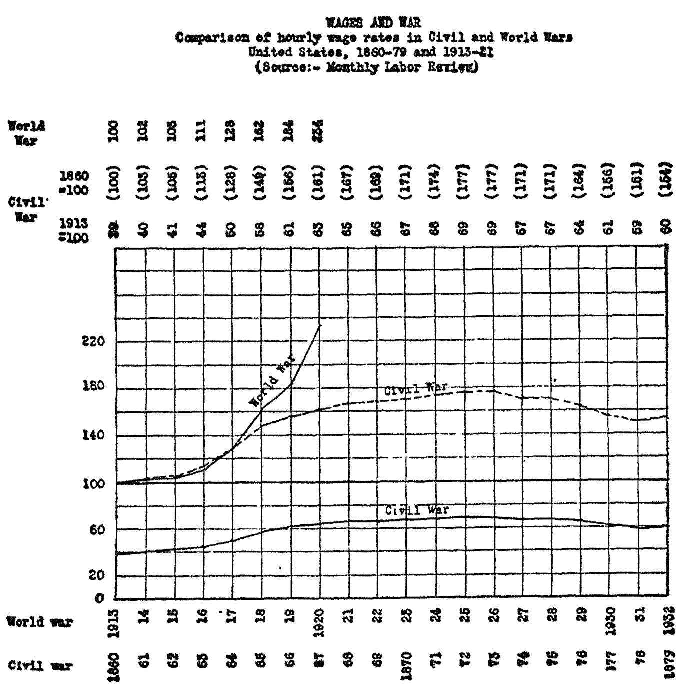


Fig. 260.

to get a common index, an average change in the cost of all foodstuffs, which could be used as a cost-of-living index figure.

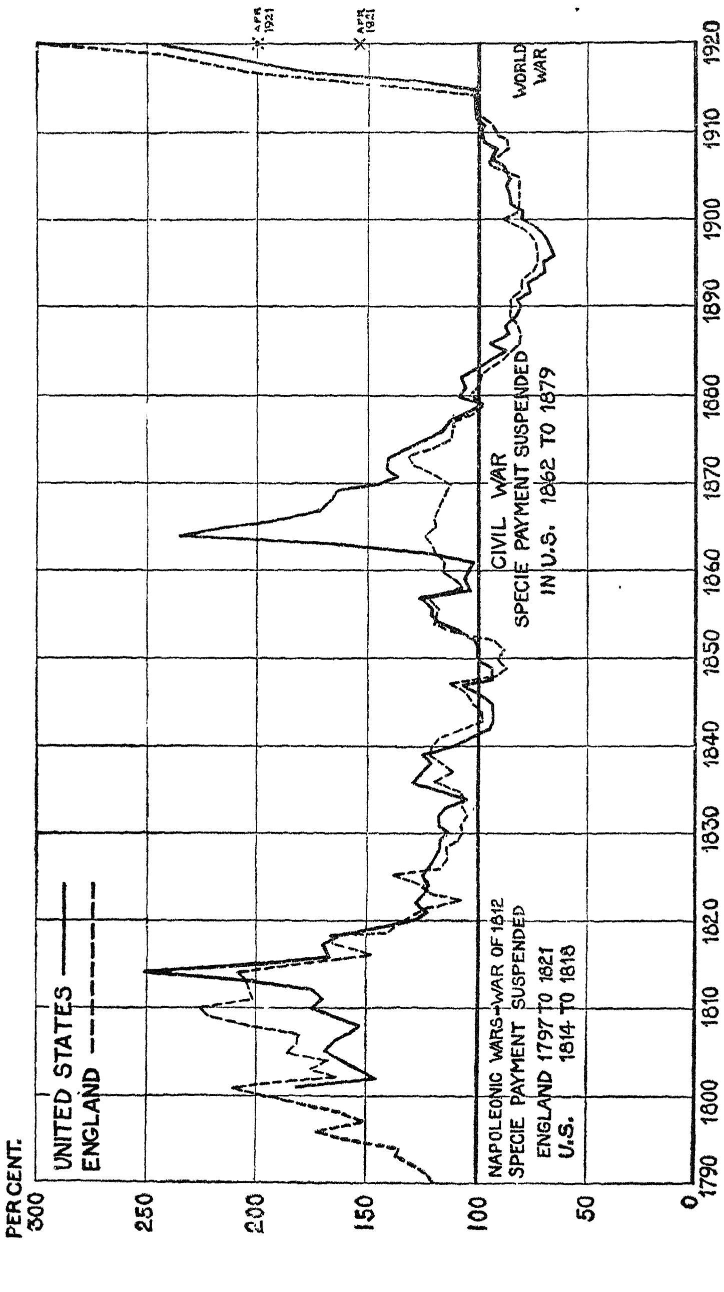


Fig. 261.

Wholesale commodity prices in the United States and England expressed as percentages of figures for 1913.—*Permission of Mr. Carl Snyder.*

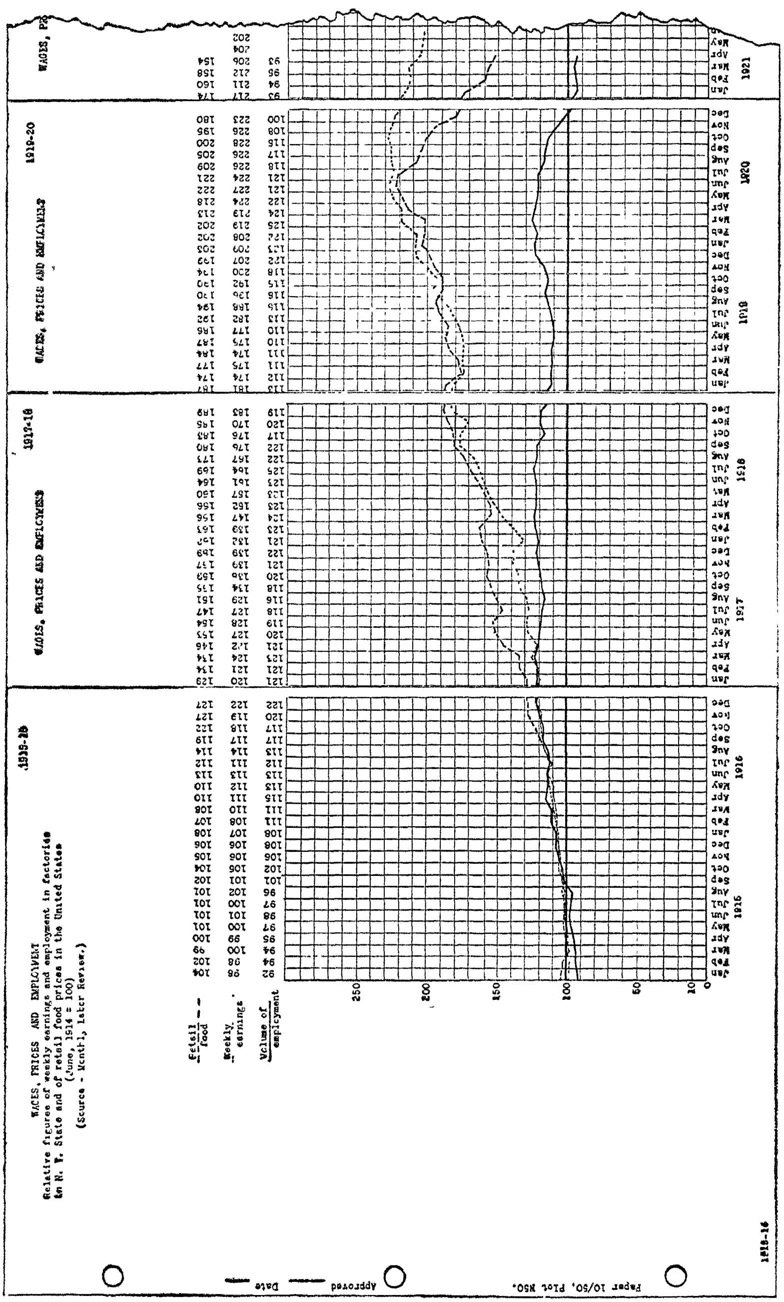

WAGES, PRICES AND EMPLOYMENT
Relative figures of weekly earnings and employment in factories in N. Y. State and of retail food prices in the United States
(June, 1914 = 100)
(Source - Monthly Labor Review.)
Retail food
Weekly earnings
Volume of employment

Fig. 262.

In this writing, nothing more can be done than to indicate the problem. It has not yet been finally answered.

As to the plotting and other details of chart-making for index numbers, the rules laid down for historical charts in general apply. The use of index numbers, however, generally brings all data to a common scale and range of variation so that a uniform charting field should be used for these charts, when you are preparing a series of them. The uniformity is of

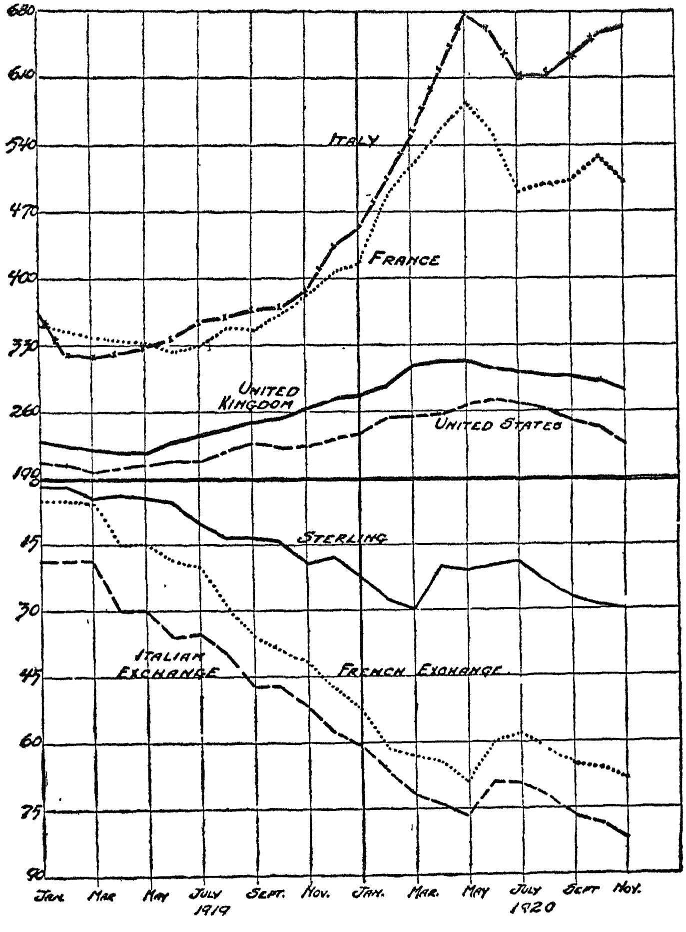


Fig. 263. Correlation Shown by Mirroring.

The chart shows Foreign Exchanges on New York below the heavy line (in terms of depreciation from parity), and commodity prices above the heavy line.—*Permission of Mr. Carl Snyder.*

value for making comparisons. It is well to place the data of the original or absolute series beside the data of index or relative figures in the data attached to the chart, whenever the relative has been computed directly from absolute data. Of course this is not desirable where the indices have been compiled from a large number of original series.

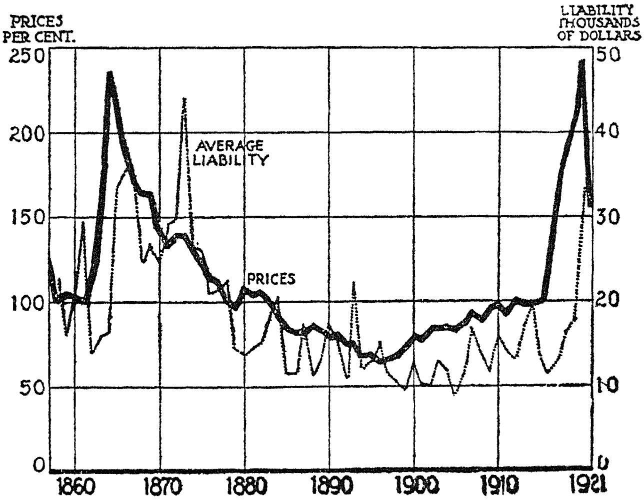


Fig. 264. A Slightly Lagged Correlation.

Average liability of failures in the United States each year compared with changes in wholesale commodity prices. (Department of Labor index.)—*Permission of Mr. Carl Snyder.*

Relative figures and index figures[2] can be frequently used for all sorts of data other than historical data, but the principles and applications are the same and the greatest use occurs for relatives and indices in historical series. They afford a simple

[2] The distinction between relative figures and index numbers is really very clear and should be adhered to. Relative numbers are directly related to absolute data; the absolute data is that original series of actual figures which has an individual as well as a collective meaning. Thus, price-quotations are absolute figures. From these absolute figures we derive relative figures by the process of division, using a constant divisor which we call the "base-figure." But the relative figures, so derived, have no individual significance; they take on a meaning only collectively, as a series, each being a ratio between two absolute figures. Relative figures are but one step removed from absolute figures.

Very different from these, are index numbers. These have no corresponding absolute data; they are merely indicators of some theoretical and wholly imaginary idea, such as the combined movement of many individual things. They are usually derived from relative figures, as explained in the text, either by simple or by weighted averaging.

and sound means of comparing any series, bringing widely different series, or even series measured in different units, together into easily compared curves. They are at bottom no more

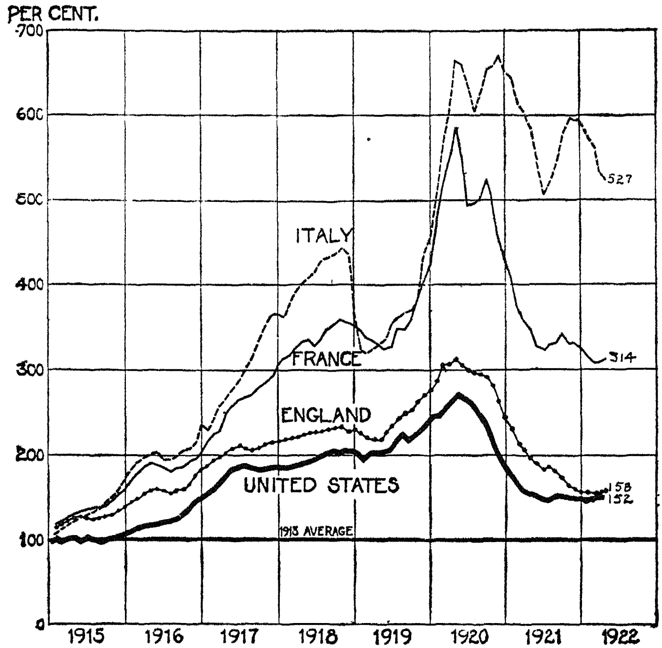


Fig. 265.
Wholesale commodity prices in four countries (average prices in 1913 = 100%). —*Permission of Mr. Carl Snyder.*

than percentages, though different from the percentages used in 100% bars and band-charts, in that the value of 100% is no longer a total, but merely one of the values in the series. That the comparison of curves is largely incidental to the search for correlation between the phenomena which the curves represent, and that correlation studies are in historical series very often directed to the practical end of forecasting or predicting future conditions, are details which in no way limit the general usefulness of indices. You may be seeking light on the probable level of prices in your business in the future; this calls for forecasting and therefore for a knowledge of attendant and preceding developments. But you may also be only interested in the relation between changes in your advertising

appropriation and in your gross sales; in this case, too, you can well use index numbers or relatives.[3]

[3] A very different type of relative figure is the "chain-percentage" or "link-relative." Being anti-logarithm of the logarithmic differential (or successive differences) of a series, it has very little value and is rather over-estimated. It is secured by taking each item as a percentage of the preceding item in a series (sometimes after subtracting 100), and has, therefore, no constant base. It is discussed in later chapters.

Chapter XXVII

FREQUENCY SERIES

We have now to consider the curves for data of a non-historical nature, that is, data in which time is not the independent variable. This is often data of conditions at a single moment of time—a cross-section, as it is sometimes called, of the phenomenon. At other times it is a compilation or recapitulation of phenomena (events or conditions) through a period of time—still, if you please, a cross-section. The analysis of such data proceeds through a series of changing conditions, and the conditions can sometimes be so coherently arranged as to form a variable. When this is the case, the data can generally be profitably shown and studied by means of a curve-chart. Curves of this nature are called "frequency curves," a name which is derived from their chief purpose, which is the display of the frequency with which the phenomena occur under given conditions. They are also sometimes called "pictograms," but the latter name has fortunately not found general acceptance.

A few examples of this type of data will serve to make the class clear. The manager of a chain of retail stores, and to a

CITY FINANCES
Percapita Revenue Receipts and Cost Payments
United States
Year Ended June 30, 1919
(Source:- United States Census)

Population of Cities (1917)	Percapita Revenue (Dollars)	Percapita Expense (Dollars)
30,000 - 50,000	27.14	26.23
50,000 - 100,000	26.23	27.29
100,000 - 300,000	29.18	32.10
300,000 - 500,000	39.53	40.16
500,000 and Over	41.87	40.78

Fig. 266.

lesser extent any distributor over a large territory, will be benefited by a report showing the per capita sales in cities of different sizes, as such statistics will show him the comparative value of large and small town outlets for his goods. Here the classification would be according to the population of towns, and for certain purposes a simple analysis would show the average per capita sales in towns of each size. Again a manufacturer is putting up his products in many different sizes and

SIZE	(48 - can) CASES
1/2 pound cans	1,200,034
1 pound cans	13,901,592
1-1/2 pound cans	1,529
2 pound cans	3,003

PRODUCTION OF RED SALMON
Output of Canned Red or Sockeye Salmon
Alaska Fisheries
7-year Total, 1913-1919
(Source:- "United States Bureau of Fisheries")

Fig. 267.

cases, and an analysis of sales according to size might take the form of a table showing sales of each size which again might be charted in a frequency curve when the sizes form a connected mathematical series. And to take one more example, the manufacturer of building materials might find advan-

RENTS IN DENMARK
Average Yearly Rentals of Family Dwellings
Danish Cities
1918 and 1919
(1 Crown at par = 26.8 cents)
(Source-- Monthly Labor Review)

	Capital		Provinces	
	1918	1919	1918	1919
	(--------Crowns-----			---)
1 room and kitchen	138	148	85	99
2 rooms and kitchen	290	304	179	198
3 rooms and kitchen	415	434	271	301
4 rooms and kitchen	544	566	379	420
5 rooms and kitchen	780	828	502	557
6 rooms and kitchen	1065	1117	632	706
7 rooms and kitchen	1369	1464	706	851
8 rooms and more	2103	2328	1013	1151

Fig. 268.

tageous a curve showing the number of one-story, two-story and higher buildings in his territory.

Both the historical and the frequency series are numerical distributions, that is, their independent variables are mathematical series or progressions. When the independent variable marks specific points or periods of time we call the numerical distribution a historical series. In all other cases we call it a frequency series. Nor is it possible to apply this distinction always, for there is a large class of numerical distributions in which time is counted—not from a single common origin-point of time but from various and usually unrecorded and unim-

EFFECTS OF DIPHTHERIA ANTITOXIN
Chances of Recovery due to Use of Antitoxin
On Various Days after Diphtheria is Discovered
(Source:- Kolle and Hetsch)

Day of Disease on which Antitoxin is first Administered	Percentage of Cases in which Recovery is Made
1	100
2	96
3	86
4	77
5	61
6	45

Fig. 269.

portant reference-points, and these also are to be classed as frequency series. Counts or classifications (i.e. distributions) of the population by age in years, or of mortality-rates, marriages, weights, heights, illiteracy, or the like, by ages; of orders or shipments by length of time taken to complete, or the like, are examples of frequency series which have as their basis, time.

In the consideration of frequency series, we may regress a moment to the general subject of statistical tabulation. For it is in frequency series that the greatest measure of statistical treatment is called for, not alone in the handling of the completed series, but in the preliminary work of compiling the series. And in actual practise the student will encounter a baffling heterogeneity of frequency distributions, presented by their compilers in various shapes and statistical fashions and often, unfortunately, in what he will come to recognize as

FIRE LOSSES IN THE UNITED STATES
Statistics of Losses of Property due to Fire in Larger Cities
1919
(Source:- National Board of Fire Underwriters)

State	City	Population	Number of Fires	Property Loss Total	Property Loss Percapita
Ala	Birmingham	225,000	2,685	591,207	2.63
Cal	Los Angeles	700,000	3,100	1,388,205	1.98
	Oakland	225,000	1,503	136,746	.61
	San Francisco	525,000	3,351	....	....
Col	Denver	290,000	1,538	374,214	1.29
Conn	Bridgeport	200,000	752	118,499	.59
	Hartford	140,000	651	215,695	1.54
	New Haven	175,000	874	265,717	1.52
	Waterbury	100,000	459	133,993	1.34
D.C.	Washington	400,000	1,562	583,171	1.46
Fla	Jacksonville	110,000	423	250,298	2.28
Ga	Atlanta	230,000	762	665,336	2.86
Ill	Chicago	2,816,000	17,208	7,331,023	2.60
Ind	Indianapolis	300,000	2,969	1,068,937	3.56
Ia	Des Moines	108,000	925	285,338	2.64
Kan	Kansas City	100,000	984	155,765	1.55
Ky	Louisville	265,000	919	552,204	2.08
La	New Orleans	380,000	882	548,248	1.44
Md	Baltimore	750,000	3,244	3,205,602	4.27
Mass	Boston	808,310	4,934	2,577,584	3.19
	Cambridge	112,000	584	418,353	3.73
	Fall River	130,000	414	210,631	1.62
	Lawrence	105,000	522	78,120	.74
	Lowell	118,000	943	232,103	1.96
	Lynn	104,000	631	96,099	.91
	New Bedford	120,000	643	249,917	2.08
	Springfield	130,000	812	357,947	2.75
	Worcester	190,000	1,345	246,839	1.30
Mich	Detroit	900,000	4,190	4,026,279	4.47
	Grand Rapids	145,000	1,093	757,604	5.22
Minn	Duluth	100,000	415	169,807	1.69
	Minneapolis	400,000	2,279	924,733	2.31
	St. Paul	275,000	1,299	633,140	2.30
Mo	Kansas City	320,000	3,297	1,027,052	3.21
	St. Louis	900,000	4,088	1,616,254	1.80
Neb	Omaha	205,000	1,233	293,446	1.43
N.J.	Camden	110,000	438	76,933	.70
	Elizabeth	110,000	410	99,013	.98
	Jersey City	300,000	1,159	413,563	1.38
	Newark	450,000	1,545	896,881	1.99
	Patterson	130,000	473	393,197	3.02
	Trenton	108,000	403	287,079	2.66
N.Y.	New York City	6,006,794	13,429	12,488,258	2.08
	Rochester	300,000	913	390,375	1.30
	Schenectady	106,000	316	69,696	.66
	Syracuse	160,000	567	253,527	1.58
	Yonkers	106,000	482	168,779	1.59
Ohio	Akron	175,000	835	389,819	2.23
	Cincinnati	418,022	1,468	612,742	1.47
	Cleveland	750,000	3,906	1,793,044	2.39
	Columbus	240,000	823	249,375	1.04
	Dayton	153,930	1,081	300,361	1.95
	Toledo	220,000	1,066	1,500,075	6.82
	Youngstown	130,000	745	199,518	1.53
Okla	Olkahoma City	110,000	510	402,080	3.65
Ore	Portland	326,000	944	552,831	1.70
Pa	Erie	112,000	425	100,267	.89
	Philadelphia	1,850,000	4,204	4,885,485	2.64
	Pittsburgh	600,000	2,580	1,707,007	2.84
	Reading	110,000	141	138,218	1.26
	Scranton	150,000	405	502,811	3.35
R.I.	Providence	260,000	1,766	627,611	2.41
Tenn	Memphis	166,000	1,738	650,993	3.92
	Nashville	150,000	667	406,751	2.71
Tex	Dallas	140,000	893	253,436	1.81
	Fort Worth	110,000	637	226,938	2.06
	Houston	150,000	1,003	1,010,052	6.73
	San Antonio	150,000	401	203,996	1.36
Utah	Salt Lake City	130,000	572	347,066	2.67
Va	Norfolk	150,000	773	4,084,267	27.23
	Richmond	170,000	741	134,426	.79
Wash	Seattle	380,000	2,358	762,757	2.01
	Spokane	157,626	728	334,617	2.13
	Tacoma	123,000	794	163,863	1.33
Wisc	Milwaukee	510,000	2,228	917,336	1.70

Fig. 270. The Raw Material for a Frequency Series.

various stages of compilation. While we cannot attempt to cover the subject as thoroughly as it is treated in the statistical text-books, yet we may well give it a brief survey, in order that the chart-maker may be enabled the better to construct his frequency curve.

The first stage in the preparation of a frequency curve is the simple listing or list of observations. This list is not in any sense a frequency series; it is merely the crude form, the raw material, from which the frequency series will be made. The stubs of the list are names, or numbers, which can be

Check number of man	Average daily output
4003	381
4182	380
4206	370
4215	392
4220	400
4221	414
4223	394
4200	413
4232	416
4238	307
4276	392
4282	374
4287	347
4289	406
4342	377
4350	428
4354	390
4356	398
4361	402
4370	387
4373	382
4392	411
4395	408
4318	410
4402	391
4419	407
4426	425
4452	399
4465	401

OUTPUT OF WORKERS
(Non-stereotyped operation)
(Source:- Report of P. S. Florence)

Fig. 271. Another Crude List.

called items, and the list itself is composed of numerical values or other observations which have been noted for these items.

It is possible for both stubs and observations to take the form of numbers, but still the list does not form a series. Also it is possible for both stubs and observations to be abstract, that is, not numerical. Usually the stubs are not numerical, while the observations are. The length of the list, that is, the number of items in it, indicates the total "population" or "universe" of the distribution or series which will be formed. A universe of much less than a hundred items is not likely to prove a very reliable "sampling;" as a rule the sampling should be considerably larger and detailed reliability can generally be had only in samplings which contain thousands of observations. The trustworthiness of a sampling also depends, of course, on the size of the external or unobserved universe, as well as upon bias and error in the selection of observed items or making of observations.

The second stage, that is, the first step in the conversion of this list into a series, is the rearrangement of the list in the order of magnitude of observations. This is done to facilitate

Bridgeport	$0.59	Fall River	1.62	Nashville	2.71
Oakland	.61	Duluth	1.69	Springfield, Ms.	2.75
Schenectady	.66	Portland, Ore.	1.70	Pittsburg	2.84
Camden	.70	Milwaukee	1.70	Atlanta	2.85
Lawrence	.74	St. Louis	1.80	Patterson	3.02
Richmond	.79	Dallas	1.81	Boston	3.19
Erie	.89	Dayton	1.95	Kansas City, Mo.	3.21
Lynn	.91	Lowell	1.96	Scranton	3.35
Elizabeth	.98	Los Angeles	1.98	Indianapolis	3.56
Columbus	1.04	Newark	1.99	Oklahoma City	3.65
Reading	1.26	Seattle	2.01	Cambridge	3.73
Denver	1.29	Fort Worth	2.06	Memphis	3.92
Worcester	1.30	Louisville	2.08	Baltimore	4.27
Rochester	1.30	New Bedford	2.08	Detroit	4.47
Tacoma	1.33	New York City	2.08	Grand Rapids	5.22
Waterbury	1.34	Spokane	2.13	Houston	6.73
San Antonio	1.36	Akron	2.23	Toledo	6.82
Jersey City	1.38	Jacksonville	2.28	Norfolk	27.23
Omaha	1.43	St. Paul	2.30		
New Orleans	1.44	Minneapolis	2.31		
Washington	1.46	Cleveland	2.39		
Cincinnati	1.47	Providence	2.41		
New Haven	1.52	Chicago	2.60		
Youngstown	1.53	Birmingham	2.63		
Hartford	1.54	Des Moines	2.64		
Kansas City	1.55	Philadelphia	2.64		
Syracuse	1.58	Trenton	2.66		
Yonkers	1.59	Salt Lake City	2.67		

PERCAPITA FIRE LOSSES
in 74 large American cities
1919

Fig. 272. The First Step is Arrangement by Magnitude.

a count, which will shortly take place, or to enable us to sum up two or more sets of observations about the same items in the list. Notice that we now arrange by the observations and not by the stubs or items. Already the emphasis has shifted

to what in a broad sense might be called in this crude list a dependent variable. The reason is that we are about to forget the stubs altogether and make the observations (which occupy the place of a dependent variable) the independent variable of our series. In other words we are about to distribute the data according to the numerical size of its parts.

The third stage, and the step which finally yields us a frequency series, is to gather the observations into groups, or

$10,000	3,400 (5)	2,320
8,000	3,350 (3)	2,300 (39)
6,250 (2)	3,300 (69)	2,275 (3)
6,000 (8)	3,250 (54)	2,270 (3)
5,750 (3)	3,200 (35)	2,250 (22)
5,500 (32)	3,150 (6)	2,220
5,250 (11)	3,100 (27)	2,200 (87)
5,000 (82)	3,060	2,100 (52)
4,850 (3)	3,050	2,000 (91)
4,800 (3)	3,000 (319)	1,900 (9)
4,750 (12)	2,900 (16)	1,800 (46)
4,600 (2)	2,870 (3)	1,750 (10)
4,500 (95)	2,850 (10)	1,700 (15)
4,400 (16)	2,820 (2)	1,650 (2)
4,350 (16)	2,800 (74)	1,600 (8)
4,200 (3)	2,750 (52)	1,500 (10)
4,150 (15)	2,700 (120)	1,467
4,000 (207)	2,690 (4)	1,400
3,937	2,650 (20)	1,250 (4)
3,900 (2)	2,640 (3)	1,000 (3)
3,850 (2)	2,600 (101)	600
3,800 (7)	2,570	300
3,781 (2)	2,520	
3,750 (49)	2,500 (197)	
3,700 (3)	2,460 (2)	
3,650 (2)	2,450	
3,600 (108)	2,420 (2)	
3,500 (117)	2,400 (91)	
3,450 (11)	2,350	
3,437 (2)	2,340 (3)	

COLLEGE PROFESSORS' SALARIES
Salaries paid to full professors in the colleges and universities (public institutions only), United States, 1920.
(Source:- U.S. Bureau of Education.)
(Total number = 2,460)

Fig. 273. A Common Tendency to Bunch Up.

classes, and record the count of the number of observations in each class. Now the observations have become the stubs and are the independent variable, while a new dependent variable has been created by the counts of the observations of each magnitude (that is, the number of observations occurring in each class). Here we have clearly an arbitrary choice of the independent variable. Had we recorded a different feature of the same original items in our crude list, we should have had a different independent variable for our final frequency distribution. Had we recorded two sets of observations for each item we should have had to make our choice between two

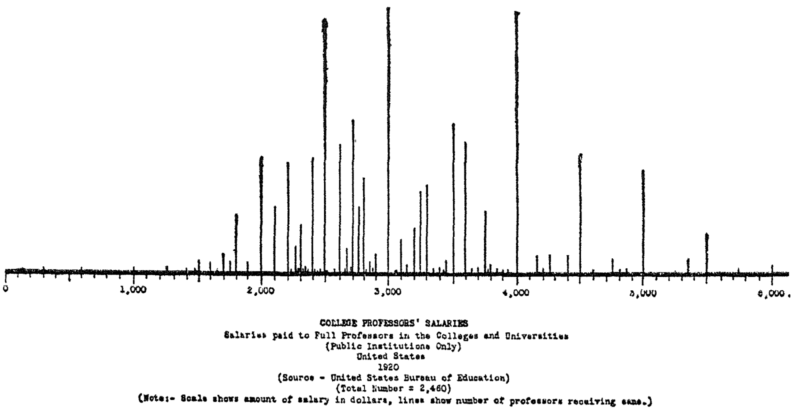


Fig. 274. Piling Up on the Round Numbers.

possible independent variables for the same series. When these alternative possibilities are presented, a wide variety of resulting series may be formed. The final dependent variable may be a count (which forms the frequency series in the strict sense) or may take the form of rates, ratios, percentages, or averages (which are only in a general sense called frequency series). It is not our purpose here to make an exhaustive study of these possible varieties and combinations; we present the reader only with a brief explanation of the simple frequency series (strictly so-called) in which the dependent variable is a count of the number of items falling within the class or group limits.

It may seem at first tnought a very simple proceeding to gather the items into classes or groups, as above described, but the fact is that at this point much statistical skill is called for.[1] For the size of the groups will determine their number, and for the best results graphically, there should be from fifteen to twenty groups. But we must not only strive for a sufficient number of groups, but we must also consider the precise location of their limits. The limits of the groups affect both their uniformity of size and their internal distributions. The last consideration is fully treated in the statistical authorities; in general, the best location of the limits from this point of view is one which places the largest number of the observations

[1] *Cf.* Yule, G. Udney, *An Introduction to the Theory of Statistics*, pp. 79–83; King, Willford I., *Elements of Statistical Method*, pp. 105–106; also Bowley, A. L., *Elements of Statistics*, and Secrist, Horace, *An Introduction to Statistical Methods.*

COLLEGE PROFESSORS' SALARIES

Salaries of Full Professors in Colleges and Universities
(in public institutions)
United States
1920
(Source: U. S. Bureau of Education)

(Total number of professors, 2,460. Average salary -- arithm mean -- $ 3,126)

$100 Classes

A Classes	B Total
Under 251	-
251-350	1
351-450	-
451-550	-
551-650	1
651-750	-
751-850	-
851-950	-
951-1,050	2
1,051-1,150	-
1,151-1,250	4
1,251-1,350	6
1,351-1,450	1
1,451-1,550	11
1,551-1,650	10
1,651-1,750	25
1,751-1,850	46
1,851-1,950	9
1,951-2,050	91
2,051-2,150	52
2,151-2,250	110
2,251-2,350	48
2,351-2,450	94
2,451-2,550	200
2,551-2,650	125
2,651-2,750	176
2,751-2,850	86
2,851-2,950	19
2,951-3,050	320
3,051-3,150	34
3,151-3,250	99
3,251-3,350	72
3,351-3,450	18
3,451-3,550	117
3,551-3,650	105
3,651-3,750	52
3,751-3,850	11
3,851-3,950	3
3,951-4,050	207
4,051-4,150	16
4,151-4,250	19
4,251-4,350	-
4,351-4,450	18
4,451-4,550	95
4,551-4,650	2
4,651-4,750	12
4,751-4,850	5
4,851-4,950	-
4,951-5,050	82
5,051-5,150	-
5,151-5,250	-
5,251-5,350	11
5,351-5,450	-
5,451-5,550	32
5,551-5,650	-
5,651-5,750	3
5,751-5,850	-
5,851-5,950	-
5,951-6,050	8
6,051-6,150	-
6,151-6,250	2
6,251-6,350	-
6,351-6,450	-
6,451-6,550	-
......	..
......	..
......	..
......	..
8,000	1
......	..
......	..
......	..
......	..
10,000	1
over	-

Column	Classes	Heading	Values (top to bottom)
C	$200 Classes	Moving totals	1, 1, -, 1, 1, -, -, 2, 2, 4, 10, 7, 12, 21, 35, 71, 55, 100, 143, 162, 158, 142, 294, 325, 301, 262, 105, 339, 354, 133, 171, 90, 135, 222, 157, 63, 14, 210, 222, 34, 19, 16, 111, 97, 14, 18, 6, 82, 82, -, 11, 11, 32, 32, 3, 3, -, 8, 8, 2, 2, -, -
D	$200 Classes	Totals, Under 351	1, -, 1, -, 2, 10, 12, 35, 55, 143, 158, 294, 301, 105, 354, 171, 135, 157, 14, 222, 19, 111, 14, 6, 82, 11, 32, 3, -, 8, 2, -
E	$200 Classes	Totals, Under 251	-, 1, 1, -, 2, 4, 7, 21, 71, 100, 162, 142, 325, 262, 339, 133, 90, 222, 63, 210, 34, 16, 97, 18, 82, -, 11, 32, 3, 8, 2, -
F	$300 Classes	Moving Totals	1, 1, 1, 1, 1, 1, -, 2, 2, 6, 10, 11, 18, 22, 46, 81, 80, 146, 152, 253, 210, 252, 342, 419, 501, 387, 281, 425, 373, 453, 205, 189, 207, 240, 274, 188, 66, 221, 225, 241, 34, 35, 111, 113, 109, 20, 18, 88, 82, 82, 11, 11, 43, 32, 35, 3, 3, 8, 8, 10, 2, 2, -
G	$300 Classes	Totals, Under 451	1, 1, 2, 10, 22, 80, 253, 342, 387, 373, 169, 274, 221, 34, 113, 18, 82, 43, 3, 8, 2
H	$300 Classes	Totals, Under 251	-, 1, 1, 2, 11, 46, 146, 210, 419, 281, 453, 207, 168, 226, 35, 109, 88, 11, 32, 3, 10, -
J	$300 Classes	Totals, Under 351	1, 1, -, 6, 18, 81, 152, 252, 501, 425, 205, 240, 66, 241, 111, 20, 82, 11, 35, 8, 2, -
K	$400 Classes	Moving totals	1, 2, 1, 1, 1, 2, 2, 6, 12, 11, 22, 28, 47, 92, 90, 171, 198, 262, 301, 304, 452, 467, 595, 587, 406, 601, 459, 472, 525, 223, 306, 312, 292, 285, 171, 273, 236, 244, 241, 50, 130, 113, 125, 115, 20, 100, 88, 82, 93, 11, 43, 43, 35, 35, 3, 11, 8, 10, 10, 2, 2, -
L	$400 Classes	Totals, Under 551	1, 1, 12, 47, 198, 452, 405, 525, 292, 236, 130, 20, 93, 35, 8, 2
M	$400 Classes	Totals, Under 251	-, 2, 2, 11, 92, 262, 467, 601, 223, 285, 244, 113, 100, 11, 35, 10, -
N	$400 Classes	Totals, Under 351	1, 1, 2, 22, 90, 301, 595, 459, 306, 171, 241, 125, 88, 43, 3, 10
P	$400 Classes	Totals, Under 451	1, 1, 6, 28, 171, 304, 587, 472, 312, 273, 50, 115, 82, 43, 11, 2
Q	$500 Classes	Moving Totals	2, 2, 1, 1, 3, 2, 6, 12, 13, 22, 32, 53, 93, 101, 181, 223, 308, 310, 395, 504, 577, 643, 681, 606, 726, 635, 558, 544, 543, 340, 423, 364, 303, 288, 378, 288, 254, 244, 257, 145, 132, 125, 131, 115, 102, 100, 88, 99, 93, 43, 43, 46, 35, 35, 11, 11, 10, 10, 10, 2, 2, -
R	$500 Classes	Totals, Under 661	2, 2, 32, 223, 577, 635, 423, 288, 132, 100, 43, 11, 2
S	$500 Classes	Totals, Under 251	-, 2, 6, 53, 308, 643, 558, 364, 254, 125, 88, 46, 10, -
T	$500 Classes	Totals, Under 351	1, 1, 12, 93, 310, 681, 544, 303, 244, 131, 99, 35, 10
U	$500 Classes	Totals, Under 451	1, 1, 13, 101, 395, 606, 543, 288, 257, 115, 93, 35, 10
W	$500 Classes	Totals, Under 551	1, 3, 22, 181, 504, 726, 340, 378, 145, 102, 43, 11, 2
Z		Class Limits	250, 350, 450, 550, 650, 750, 850, 950, 1,050, 1,150, 1,250, 1,350, 1,450, 1,550, 1,650, 1,750, 1,850, 1,950, 2,050, 2,150, 2,250, 2,350, 2,450, 2,550, 2,650, 2,750, 2,850, 2,950, 3,050, 3,150, 3,250, 3,350, 3,450, 3,550, 3,650, 3,750, 3,850, 3,950, 4,050, 4,150, 4,250, 4,350, 4,450, 4,550, 4,650, 4,750, 4,850, 4,950, 5,050, 5,150, 5,250, 5,350, 5,450, 5,550, 5,650, 5,750, 5,850, 5,950, 6,050, 6,150, 6,250, 6,350, 6,450, 6,550

Fig. 275. Comparison of Fourteen Series Derived from the Same Data by the Use of Different Group Limits and Group Sizes.

which belong to the group in or near the center of the group.[2] Thus if we are counting men of various heights, and notice a

[2] This makes each group include all observations of doubtful accuracy, such as the observations at and immediately about the round numbers.

COLLEGE PROFESSORS SALARIES
Salaries of Full Professors in Colleges and Universities
(In Public Institutions)
United States
1920
(Source.- United States Bureau of Education)
(Note.- All data is for $300-groups - Series G from $451 - 750, 751 - 1,050, etc.;
Series H from $251 - 550, 551 - 850, etc.; - Series J from $351 - 650, 651 - 950, etc.;
the total of each series being the same, 2,460.)

Series J —— Salary	Number	Series H - - - Salary	Number	Series G ······· Salary	Number
	1	$ 400	1		1
$ 500	1	700	1	$ 600	1
800	•	1,000	2	900	2
1,100	6	1,300	11	1,200	10
1,400	13	1,600	46	1,500	22
1,700	81	1,900	146	1,800	80
2,000	152	2,200	210	2,100	253
2,300	252	2,500	419	2,400	342
2,600	501	2,800	281	2,700	387
2,900	425	3,100	453	3,000	373
3,200	205	3,400	207	3,300	189
3,500	240	3,700	168	3,600	274
3,800	66	4,000	225	3,900	221
4,100	241	4,300	35	4,200	34
4,400	111	4,600	109	4,500	113
4,700	20	4,900	88	4,800	18
5,000	82	5,200	11	5,100	82
5,300	11	5,500	32	5,400	43
5,600	35	5,800	5	5,700	5
5,900	6	6,100	10	6,000	8
6,200	2			6,300	2
Over	2	Over	2	Over	2

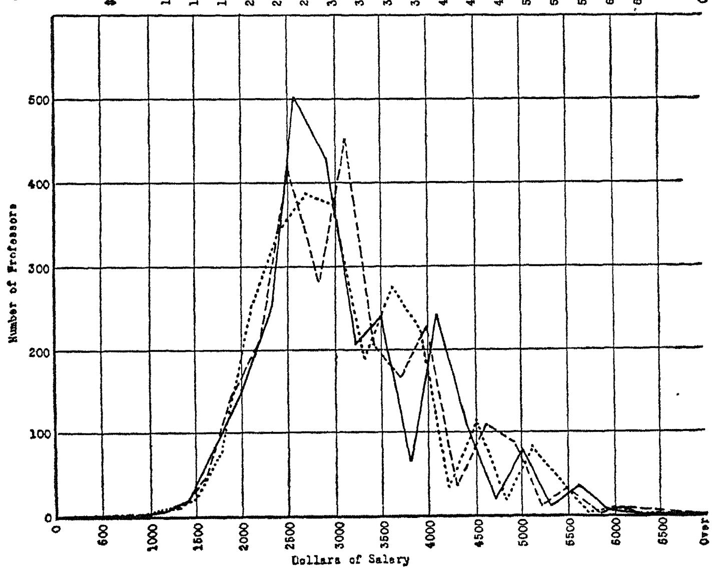


Fig. 276. Comparison of Curves of Three Series Derived from the Same Data.

tendency of the records to bunch up heavily at the round numbers (which is only natural in such measurements) we should do well to make our groups run from half inch to half inch, so that each full inch will be in the center of its group.

Strictly speaking, the ordered list may be considered a frequency series with such minute groups that there is but one stub-value (however frequent it be) in each group; that is, that

each group contains but one value of the independent variable. The series is generally unsatisfactory, however, because of its unwieldy length and its many omitted groups or classes (that is, classes with zero frequencies). We are therefore called upon to make larger groups that they may be fewer in number. As we increase the size and reduce the number of these groups, we find the curve becoming more smooth in outline, the zero-frequency groups disappearing. When carried out in detail, the process is very like the moving-total operation which we have seen performed on historical series, the same smoothing out of insignificant wrinkles being the result. However, unlike the historical series, there is no natural cyclic period to guide us in determining the lengths of final intervals. Hence when we have found the smoothest intervals, we shall take out only the totals (not the moving totals) for publication. If the results are to be published to the layman it is well to adopt round number intervals or class-limits, for his convenience, however much the data may tend to "bunch up," as previously mentioned upon the round numbers. When the series is to be presented ,to statisticians or used in research work, and such "bunching up" is noticeable, care must be taken to select intervals or class-limits which will, so far as possible, place the round numbers near the center of each class, and the class-limits will therefore be fractions rather than round numbers.

Care must also be taken that the limits of the groups be explicitly stated so that no confusion will result in the mind of the reader. Thus it would be wrong to write "100–200, 200–500, 500–1000," etc., in a table of the sizes of cities by population. Such a series should be "100–199, 200–499, 500–999," or "101–200, 201–500, 501–999," etc., as the case may be. When fractions are present, as, for example, in a similar series of the sizes of farms by acres, the best statement is "100 and less than 200, 200 and less than 500, 500 and less than 1000," etc.; but sometimes a shorter form, such as "100–199, 200–499, 500–999," etc., will not be misunderstood. Whenever space allows and there is any doubt as to either limits or the mid-points of the range, two stub columns should be used, the first to give approximate values of mid-points of each group and the second to give intervals or group limits.

To the feature of uniformity of size of groups or classes, much importance is commonly attached by statisticians, for the convenience which will result in plotting and other analy-

sis. Obviously when a portion of the series contains groups of half-inch size or range and other portions contain groups of whole inch size or range, the two kinds of groups are not di-

PERCAPITA FIRE LOSSES
in 74 large American cities
1919

Percapita Fire Loss	Number of cities
Under $.50	--
$0.51-0.75	5
0.76-1.00	4
1.01-1.25	1
1.26-1.50	12
1.51-1.75	10
1.76-2.00	6
2.01-2.25	7
2.26-2.50	5
2.51-2.75	6
2.76-3.00	2
3.01-3.25	3
3.26-3.50	1
3.51-3.75	3
3.76-4.00	1
4.01-4.25	-
4.26-4.50	2
4.51-4.75	-
4.76-5.00	-
5.01-5.25	1
5.26-5.50	-
5.51-5.75	-
5.76-6.00	-
6.01-6.25	-
6.26-6.50	-
6.51-6.75	1
6.76-7.00	1
Over 7.00	1

Fig. 277. The Frequency Series.

rectly commensurable and comparable. It is therefore always a relief to discover that the compiler of a frequency series has been able to adopt groups or classes with regular intervals between their limits. The fact remains, however, that with a very large proportion of business and sociological data the uniform group distribution is neither convenient nor satisfactory. There are cases in which the entire range, as it is called, of the distribution or series, is very great, and the great mass of observations occur near one end.[3] To show the nature of the

[3] "The general rule that intervals should be equal must not be held to bar the analysis by smaller equal intervals of some portion of the range over which the frequency curve varies very rapidly."—Yule, G. Udney, *An Introduction to the Theory of Statistics*, p. 83.

distribution through this densely "populated" portion of the range, small groups or intervals must be adopted; but to prevent an excessively long and tedious, often fruitless detail in the remainder of the series, larger intervals must be used in the sparse portions of the distribution. In such cases we are obliged to alter the sizes of the intervals. Sometimes the

DURATION OF STRIKES
United States
1921
(Source:- Monthly Labor Review)

Days of Duration		Number of Strikes
Approximate	Range	
1/4	0 - 1/2	32
1	1/2 - 1-1/2	25
2	1-1/2 - 2-1/2	42
3	2-1/2 - 3-1/2	43
4	3-1/2 - 4-1/2	43
5	4-1/2 - 5-1/2	32
6	5-1/2 - 6-1/2	32
7	6-1/2 - 7-1/2	41
8	7-1/2 - 8-1/2	27
9	8-1/2 - 9-1/2	18
10	9-1/2 - 10-1/2	40
11	10-1/2 - 11-1/2	18
12	11-1/2 - 12-1/2	11
13	12-1/2 - 13-1/2	14
14	13-1/2 - 14-1/2	24
15 - 18	14-1/2 - 18-1/2	69
19 - 21	18-1/2 - 21-1/2	42
22 - 24	21-1/2 - 24-1/2	16
25 - 28	24-1/2 - 28-1/2	30
29 - 31	28-1/2 - 31-1/2	31
32 - 35	31-1/2 - 35-1/2	34
36 - 42	35-1/2 - 42-1/2	50
43 - 49	42-1/2 - 49-1/2	37
50 - 63	49-1/2 - 63-1/2	77
64 - 77	63-1/2 - 77-1/2	57
78 - 91	77-1/2 - 91-1/2	55
92 - 199	91-1/2 - 199-1/2	165
Over 200	199-1/2 and over	42
	Total	1,147

Fig. 278.

entire range is so excessively great that no two groups can be of the same size, and it is necessary that the groups increase progressively throughout the series.

In dealing with such unevenly-grouped series, the analogy of the historical series is useful. What would you do if asked to make a curve of a historical series, let us say, the world's production of gold since the voyage of Columbus, in which the data covers at first centuries, then ages, then decades, then quinquennial periods, and lastly, individual years. Clearly you

could not directly plot the points for the data and connect the points to make a curve. You would have to change the data to uniform time intervals before plotting the curve. You

SIZE OF FARMS
United States
1920
(Source:- Census)

Acreage	Number
Less than 3	20,350
3 and less than 10	268,422
10 " " " 20	507,762
20 " " " 50	1,503,734
50 " " " 100	1,474,758
100 " " " 175	1,449,658
175 " " " 260	530,795
260 " " " 500	475,692
500 " " " 1000	149,818
1000 " over	67,387
TOTAL	6,448,358

Fig. 279.

might sum up the parts of centuries into totals or even moving totals for hundreds of years and so get a curve of 100-year production. Or you could divide the earlier data so as to get the

GOLD PRODUCTION OF THE WORLD
Estimated
1493-1919
(Source:- United States Statistical Abstract)

Period Covered	Value in Dollars
1493 - 1600	501,640,000
1601 - 1700	606,315,000
1701 - 1800	1,262,805,000
1801 - 1820	194,215,000
1821 - 1840	229,320,000
1841 - 1860	1,696,909,000
1861 - 1870	1,263,015,000
1871 - 1880	1,067,559,600
1881 - 1890	1,074,950,500
1891 - 1900	2,101,240,900
1901 - 1905	1,613,098,600
1906 - 1910	2,167,604,800
1911 - 1915	2,295,859,678
1916	454,176,500
1917	419,422,100
1918	383,605,552
1919	365,166,077

Fig. 280.

average 10-year production in the earlier periods and sum up the latter portions into ten-year groups, so getting a curve of production by decades. The intervals can be chosen at what

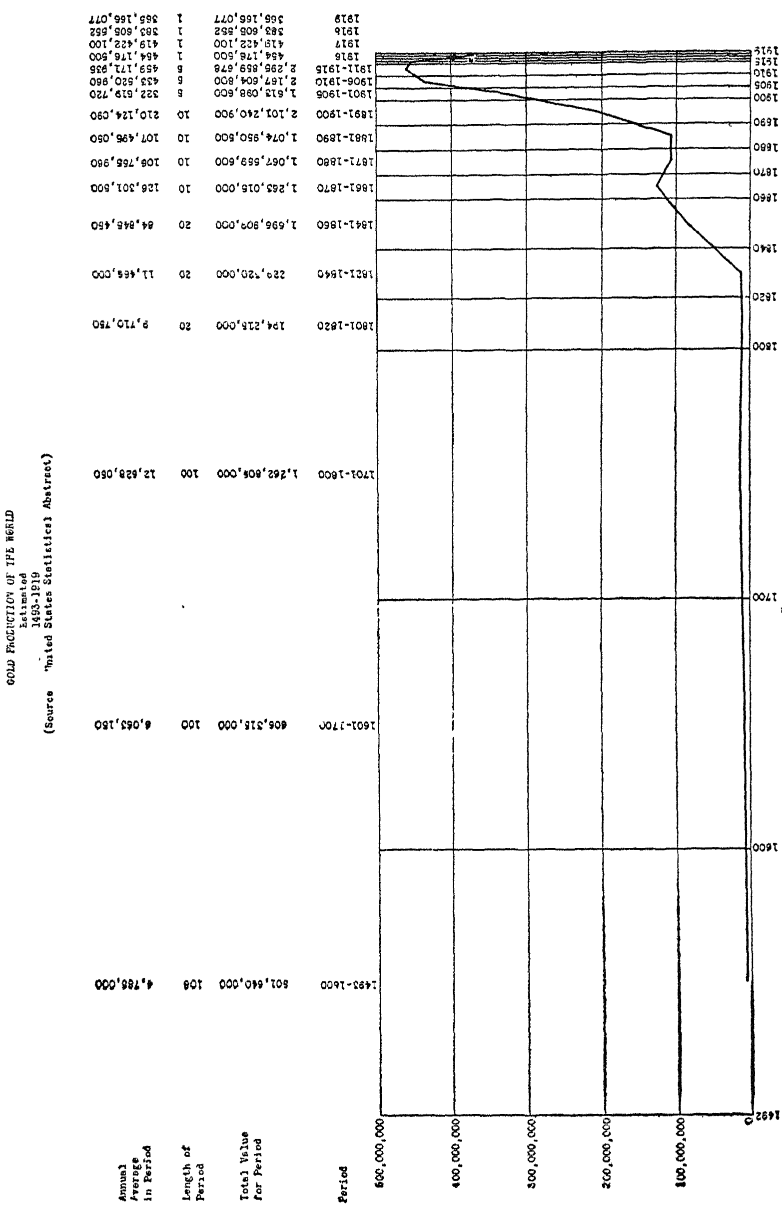


Fig. 281.

ever size you wish, the point is that you must, for the sake of the curve itself, convert the data into equivalent data for uniform intervals of time.

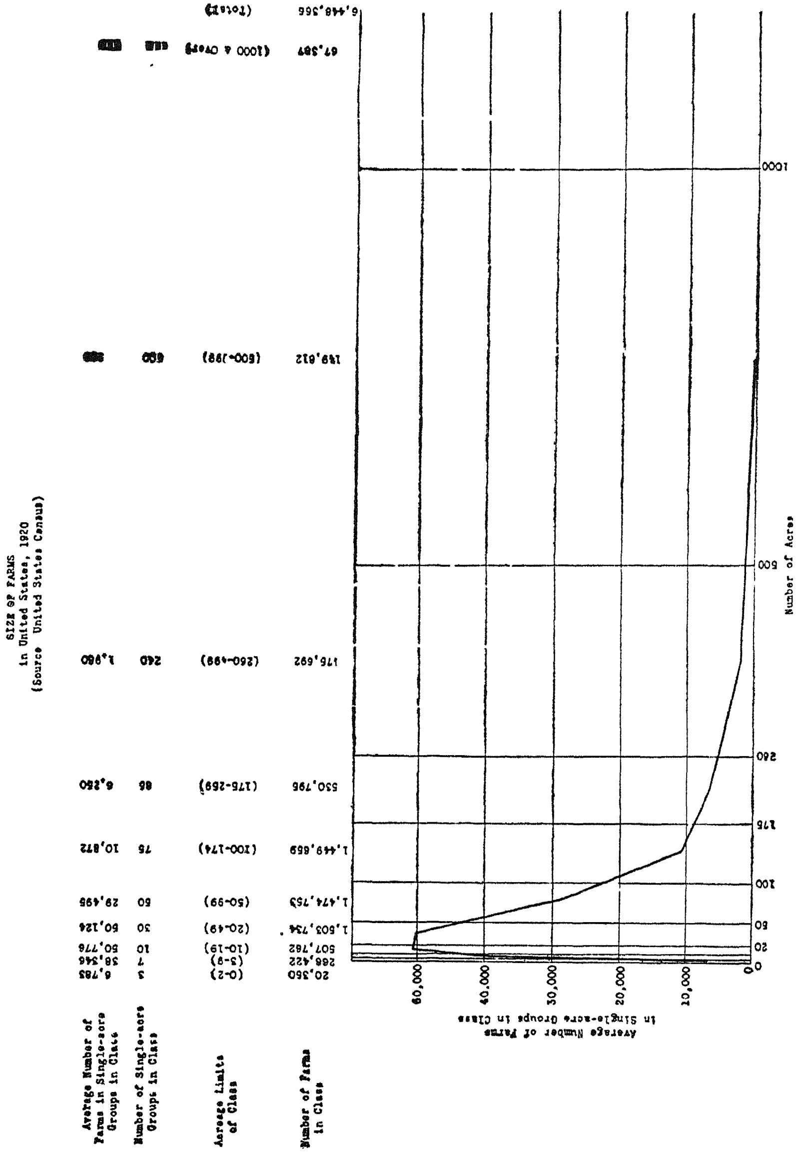


Fig. 282.

Precisely the same operation must be performed on the frequency series with irregularly-sized groups. The fact that for curves of cumulations (either of frequency or historical data) this regularity of intervals is not necessary, sometimes makes the cumulative curve, which we shall discuss in another chap-

DURATION OF STRIKES
United States
1921
(Source:- Monthly Labor Review)

Groups (Length of Strikes)	Total Frequency in Groups	Number of Equal (1-day) Sub-groups in Groups	Average Frequency in Sub-groups
0 - 1/2	32	1/2	64
1/2 - 1-1/2	25	1	25
1-1/2 - 2-1/2	42	1	42
2-1/2 - 3-1/2	43	1	43
3-1/2 - 4-1/2	43	1	43
4-1/2 - 5-1/2	32	1	32
5-1/2 - 6-1/2	32	1	32
6-1/2 - 7-1/2	41	1	41
7-1/2 - 8-1/2	27	1	27
8-1/2 - 9-1/2	18	1	18
9-1/2 - 10-1/2	40	1	40
10-1/2 - 11-1/2	18	1	18
11-1/2 - 12-1/2	11	1	11
12-1/2 - 13-1/2	14	1	14
13-1/2 - 14-1/2	24	1	24
14-1/2 - 18-1/2	69	4	17.2
18-1/2 - 21-1/2	42	3	14
21-1/2 - 24-1/2	16	3	5.33
24-1/2 - 28-1/2	30	4	7.5
28-1/2 - 31-1/2	31	3	10.3
31-1/2 - 35-1/2	34	4	8.5
35-1/2 - 42-1/2	50	7	7.14
42-1/2 - 49-1/2	37	7	5.29
49-1/2 - 63-1/2	77	14	5.50
63-1/2 - 77-1/2	57	14	4.07
77-1/2 - 91-1/2	55	14	3.93
91-1/2 - 199-1/2	165	108	1.53
Over			

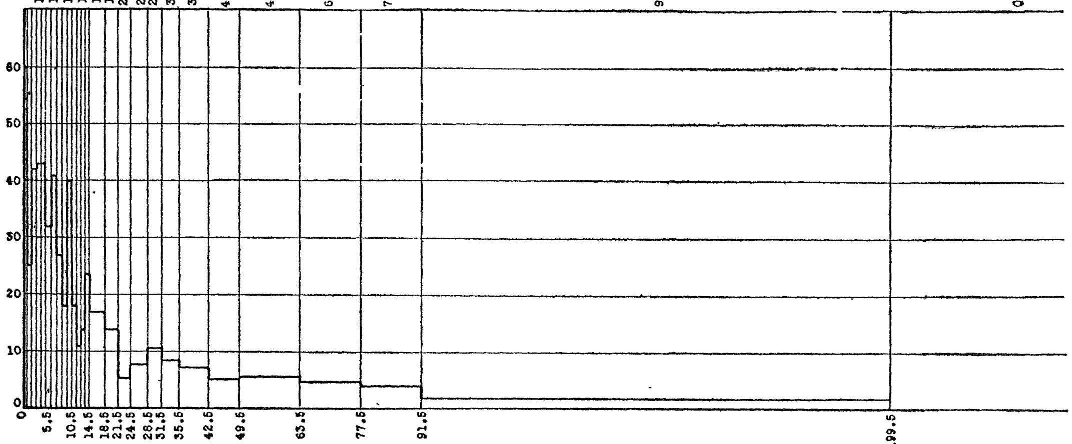


Fig. 283.

ter, far more convenient; but if we are to plot the uncumulated frequency curve, and have unequal classes or groups we must calculate the equivalents for equal intervals before we can plot the curve. And if, as often happens, a terminal class (group at one end of the series) be indeterminate, that is, have no maximum limit, so that we do not know its group range and cannot compute an equivalent figure, then we must simply omit the last group from the series, leaving the reader of the chart to conjecture that the curve runs off towards infinity.

CHAPTER XXVIII

FREQUENCY CURVES

Statisticians make a distinction between "discrete" and "continuous" frequency series, which to the chart-maker is of some assistance in determining the plotting points of the data in a frequency curve.[1] Discrete data is that in which the independent variable proceeds by leaps and bounds, lighting usually only upon the whole numbers or integers. When this last is the case, the series may be said to comprise "integral variates." Thus buildings may be classified by the number of stories or rooms they contain. Here you will find only regular intervals of one integer each, since fractional stories (barring the so-called half story) and fractional rooms can hardly be said to exist. Leaves may be classified by the number of their ribs, flowers by the number of their petals, sales-forces, departments, and establishments by the number of their employees, and cities by their populations. All of these cases are examples of discrete series.

Continuous series are those in which the phenomena may vary by infinitely small gradations, the data comprising what are called "graduated variates." Thus, of we examine the height or weight of human beings, we find them varying by the smallest possible amounts. Property classified as to value, crops as to volume in bushels, tons, or the like, farms as to area in acres, square miles, etc., and sales as to sizes, are a few examples of continuous series. In business and economics much the greater part of frequency data is of this type. And while it is not always so, yet it is ordinarily in business statistics true that continuous series require irregular group-ranges and intervals, discrete series usually falling into equal groups. The distinction between discrete and continuous data becomes less sharp in those cases of discrete data which cover large

[1] *Cf.* King, Willford I., *Elements of Statistical Method*, p. 106.

ranges, such as cities classified by populations, for here it becomes necessary to adopt arbitrary groupings which are similar to continuous data groupings. For small ranges the discrete data usually requires no arbitrary grouping together, as it automatically groups itself, and the continuous data is distinguished by the fact that group limits have to be arbitrarily set for its distribution.

MEMBERSHIP OF STRIKES
Number of Persons Involved in Strikes
United States
1921
(Source:- Monthly Labor Review)

Striking Persons Involved	Number of Strikes	Group-Range in Units of 10 Persons	Average No. in Equivalent 10-Person Class
x	y	Dx/10	y/(Dx/10)
1 - 10	219	1	219
11 - 25	286	1.5	190.6
26 - 50	252	2.5	100.8
51 - 100	214	5	42.8
101 - 250	216	15	14.4
251 - 500	153	25	6.12
501 - 1,000	101	50	2.02
1,001 - 10,000	126	900	0.14
Over 10,000	14	...	....

Fig. 284. Period Data.

For the chart-maker, the more valid distinction of frequency series is between what might be called "point-data" and "period data." The former, point-data, is that which refers to isolated, separate, and non-contiguous points along the range of the independent variable. The mortality rates at various ages are of this type and to this type belong a large class of continuous series which comprise rates, ratios, percentages, averages, and other comparisons between different basic frequency series. Most discrete series may be classed as point-data. Period data, to which most continuous series belong, is that which covers connected and conterminuous groups or classes along the range, applying throughout the groups from one group-limit or interval to the next. The student is already familiar with this distinction in the matter of historical series,

in which flow or stream figures, for example, cover periods of time and stock or fund figures, for example, refer to points of

SIZE OF FACTORIES
Manufacturing Establishments
Classified as to Number of Employees
United States
1914
(Source:- United States Census)

Number of Employees per Establishment	Establishments		Employees	
	Number	Percent	Number	Percent
0	32,856	11.9	....	....
1 - 5	140,971	51.1	317,216	4.5
6 - 20	54,379	19.7	606.594	8.6
21 - 50	22,932	8.3	742,529	10.6
51 - 100	11,079	4.0	791,726	11.3
101 - 250	8,470	3.1	1,321,077	18.8
251 - 500	3,108	1.1	1,075,108	15.3
501 - 1000	1,348	.5	926,828	13.2
Over 1,000	648	.2	1,255,259	17.8
Total	275,791	100.0	7,036,337	100.0

Fig. 285. **Period Data.**

time. Needless to say, point-data is normally plotted upon the ordinates of the chart; period data is normally plotted in the spaces between the ordinates.

VALUE OF MANUFACTURED PRODUCTS
Establishments Classified as to Value of Products, with Number of Employees in Same
United States
1914
(Source:- United States Census)

	Establishments		Employees		Value of Products	
	Number	Percent	Number	Percent	Dollars	Percent
Less than $5,000	97,061	35.2	129,623	1.8	233,381,081	1.0
$5,000 and less than $20,000	87,931	31.9	429,037	6.1	905,693,168	3.7
$20,000 and less than $100,000	56,814	20.6	999,600	14.2	2,550,229,411	10.5
$100,000 and less than $1,000,000	30,166	10.9	3.002,071	42.7	8,763,070,135	36.1
$1,000,000 and over	3,819	1.4	2,476,006	35.2	11,794,060,929	48.6
Total	275,791	100.0	7,036,337	100.0	24,246,434,724	100.0

Fig. 286. **Period Data.**

We come now to the last, and what is ostensibly the most important division of frequency curves, namely whether the

curve shall be in staircase form or smoothed. The staircase form, which really represents a collection of vertical bars, is often called a histogram, and the smoothed-curve a frequency

HOURS OF LABOR
Number of Wage Earners Employed in Manufactures
According to Prevailing Hours of Labor
United States
1914
(Source:- Statistical Abstract)

Hours of Labor	Employees	
	Number	Percent
48 and under	831,779	11.8
Between 48 and 54	944,562	13.4
54	1,813,079	25.8
Between 54 and 60	1,547,374	22.0
60	1,484,662	21.1
Between 60 and 72	249,026	3.5
72	106,080	1.5
Over 72	59,775	.8

Fig. 287. Point-and-period Data.

polygon. Of course, as we noticed in historical curves, the staircase form, if the number of steps or bars be great enough, will closely approximate the smoothed curve in appearance,

ECONOMICAL SPEEDS OF TRUCKS
Maximum Speeds at which Loaded Trucks can be Driven
Without Reducing Life of Tires

Truck Tonnage	Miles per Hour
$\frac{3}{4}$ and 1	19
$1\frac{1}{2}$	17
2	15
$2\frac{1}{2}$ - $3\frac{1}{2}$	13
4	11
5 - 7	9

Fig. 288. Point-and-period Data.

and there remains little reason to maintain it. But for series of but a few items or number of groups, the distinction between the two is great and each has its special advantages and proper uses.

The staircase curve or histogram is always more accurate for period data, in that it preserves the exact areas underneath

the curve between each set of ordinates or group-limits. The reader who recalls the belittling of area-representations for charts in which we have early indulged in this book, may find

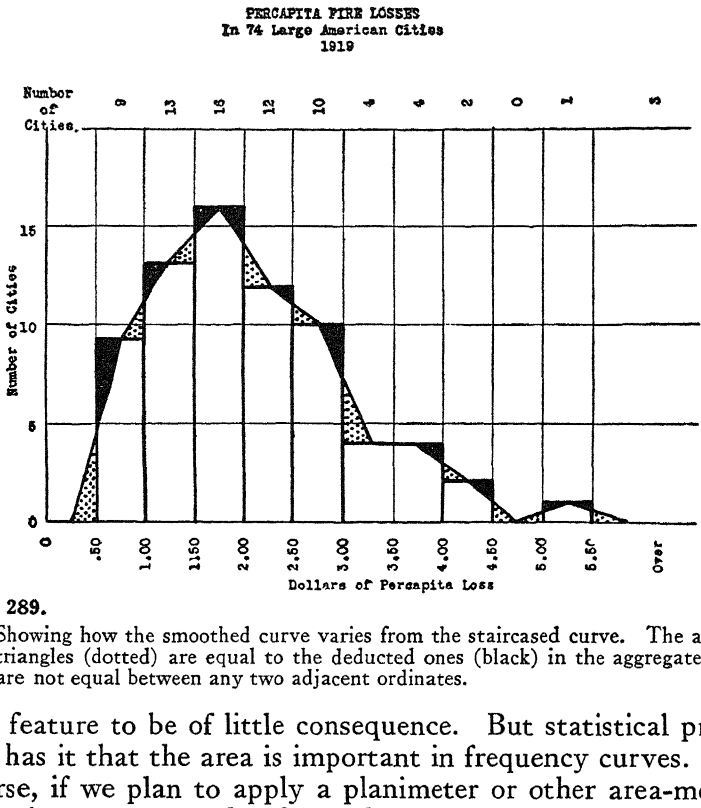


Fig. 289.

Showing how the smoothed curve varies from the staircased curve. The added triangles (dotted) are equal to the deducted ones (black) in the aggregate, but are not equal between any two adjacent ordinates.

this feature to be of little consequence. But statistical practise has it that the area is important in frequency curves. Of course, if we plan to apply a planimeter or other area-measuring instrument to the chart, the area is of real importance. Otherwise it is generally to be relegated to the limbo of academic and scientific interests. We should, however, bear it in mind, that we may the more correctly interpret our charts and base analysis upon them.

Not only is all period data more accurately represented by the individual group-areas under the staircase-curve than by the individual group-areas under a smoothed curve, but also much point-data, if we class discrete series as point-data. To be strictly accurate, discrete data should not be shown by a connected curve at all, but by separate bars; for there are no intermediate observations and the connection-line which forms the curve has no meaning over intermediate spaces on the chart.

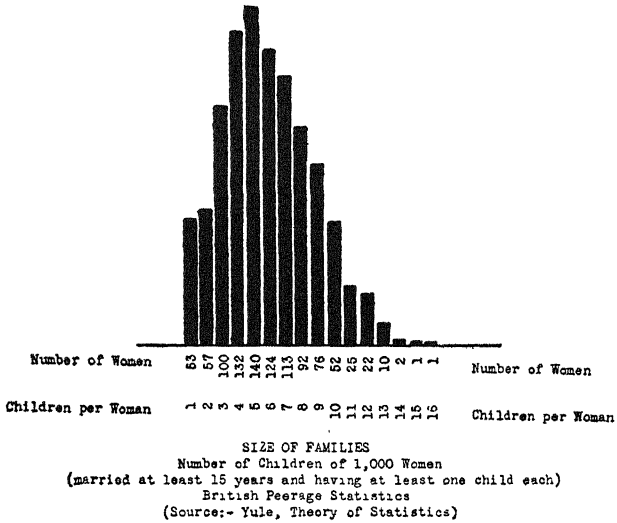
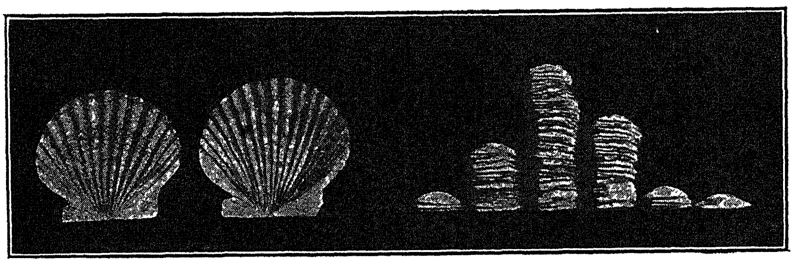


Fig. 290. The Staircased Form is Appropriate.

The chart-maker has largely to use his own judgment for plotting discrete data, as he can almost equally well, for different purposes, use the different methods of separate vertical bars, staircase or bar-like curves, and smoothed curves, not to mention plotting upon or between the ordinates.



By C. B. Davenport, Permission of Popular Science Monthly.

Fig. 291. A Very-Simplest Staircase Curve.

Showing the distribution of scallop-shells by number of ridges.

The disadvantages of the staircase form are many. In the first place, as in historical curves, it is more difficult to distinguish a number of curves brought together for comparison when they cross each other frequently. In the second place, for period data, though not for discrete data, it is less significant than the smoothed form. For while the data changes abruptly

from group to group, the phenomenon observed usually changes gradually, the values usually merging between groups. This is the more obvious if by a rearrangement of the original data we produce more and smaller groups, for then the new groups created take intermediate values. So to chart this data by a staircase curve is to give a wholly meaningless sudden change between groups, while to chart it by a smoothed curve is to bring out to the readers of the chart more clearly the gradual nature of these changes. In short the smoothed curve or frequency polygon has a truer significance than the staircase form or histogram, for period data.

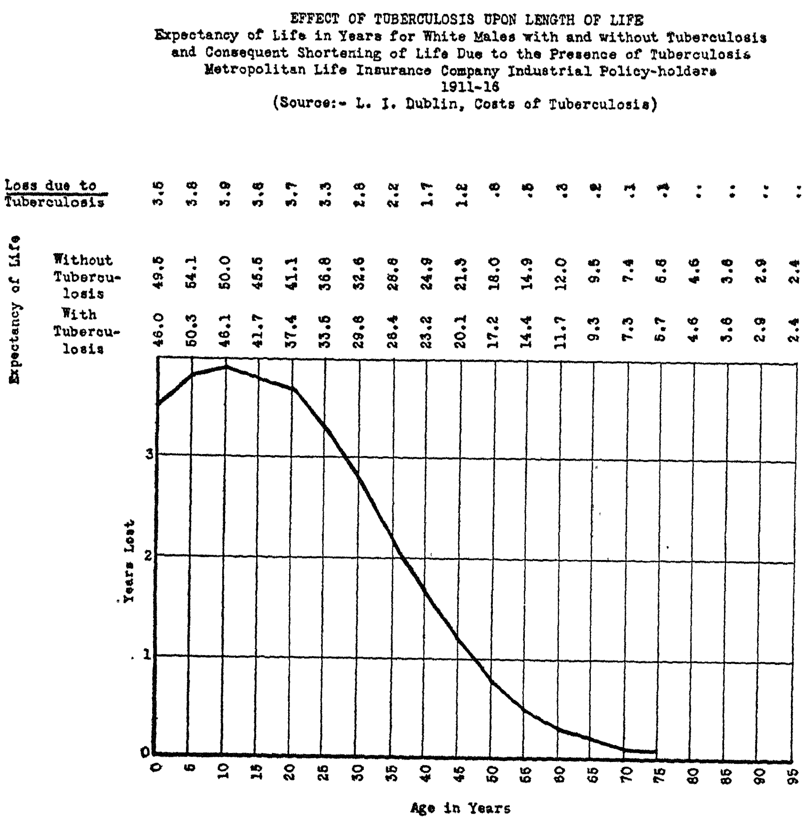


Fig. 292. The Smoothed Form is Necessary.

For continuous point-data, that is, for point-data other than discrete series, the smoothed curve is often the only possible form, the staircase form being out of the question. For

as the data represents observations at isolated points only, along the range (the x-axis scale) it is to be assumed that for intervening points intermediate values obtain, and the staircase form would be not only lacking in significance, but also in accuracy. The dependent variable in continuous point-data is usually the resultant of a process of comparison of two or more different frequency series, being ordinarily expressed as a rate, percentage, average or other ratio—a series of fractions, if you will, in which the denominators are not constant. Point data is to be found in a wide variety of forms, but is almost always in essence a derived series of this sort. The processes which yield point-data cannot be described as simply as the

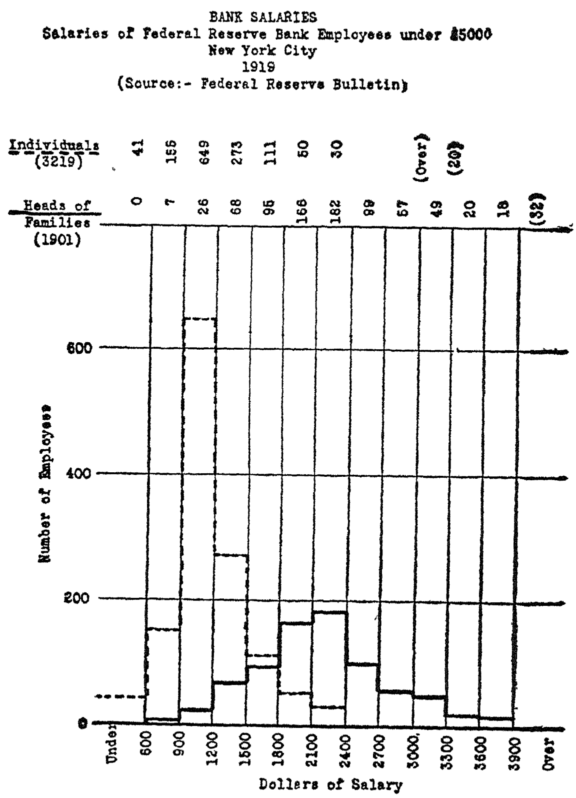


Fig. 293. It is Difficult to Compare Two Staircased Curves.

processes which yield period data; for they are also of almost unlimited variety and we shall not attempt their discussion. It is worthy of notice, however, that during such preliminary

steps in the comparison of two frequency series we commonly find the gun-shot plotting method useful.

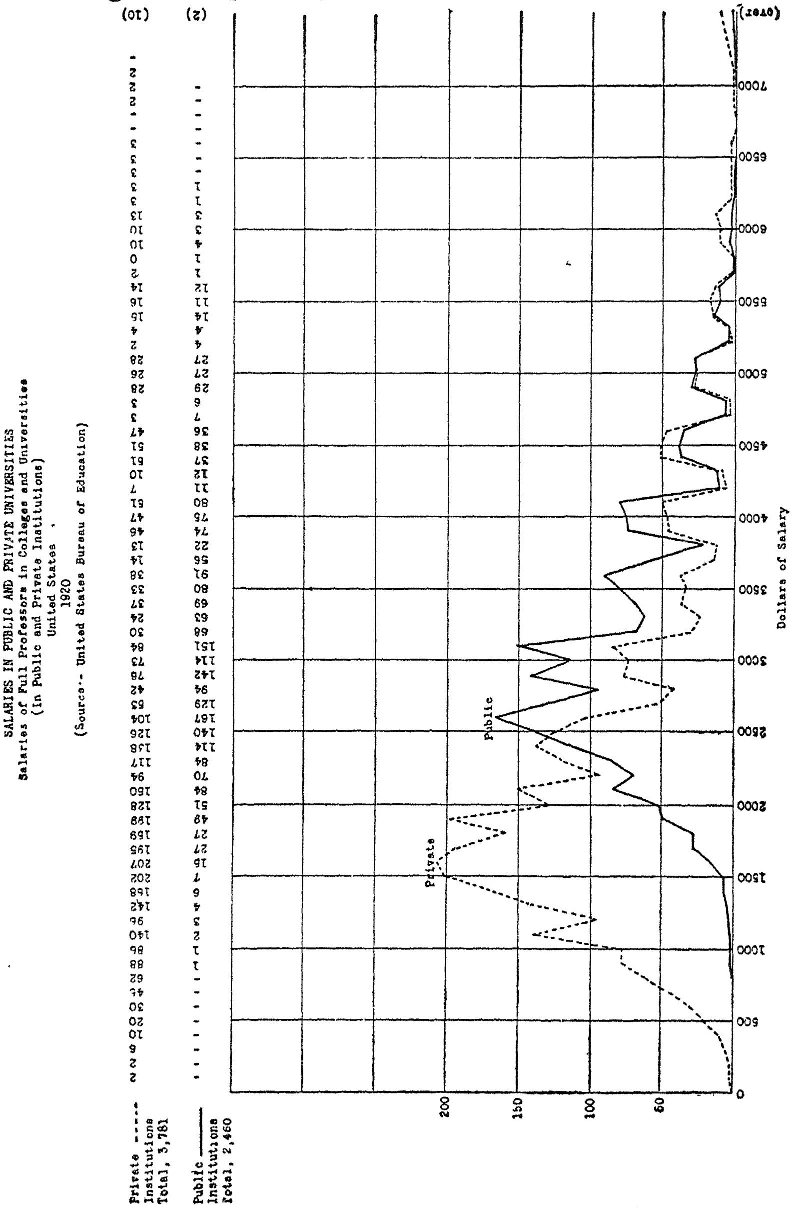


Fig. 294. A Cumulable Series.

The use of the term frequency series for point data, resulting from the comparison of two series, is permissible only

in a broad sense, the point-data series not being a distribution displaying the frequencies of the phenomena in the specified groupings. Such point data, like the balances, stocks-on-hand, or fund figures, in historical series cannot be cumulated, and

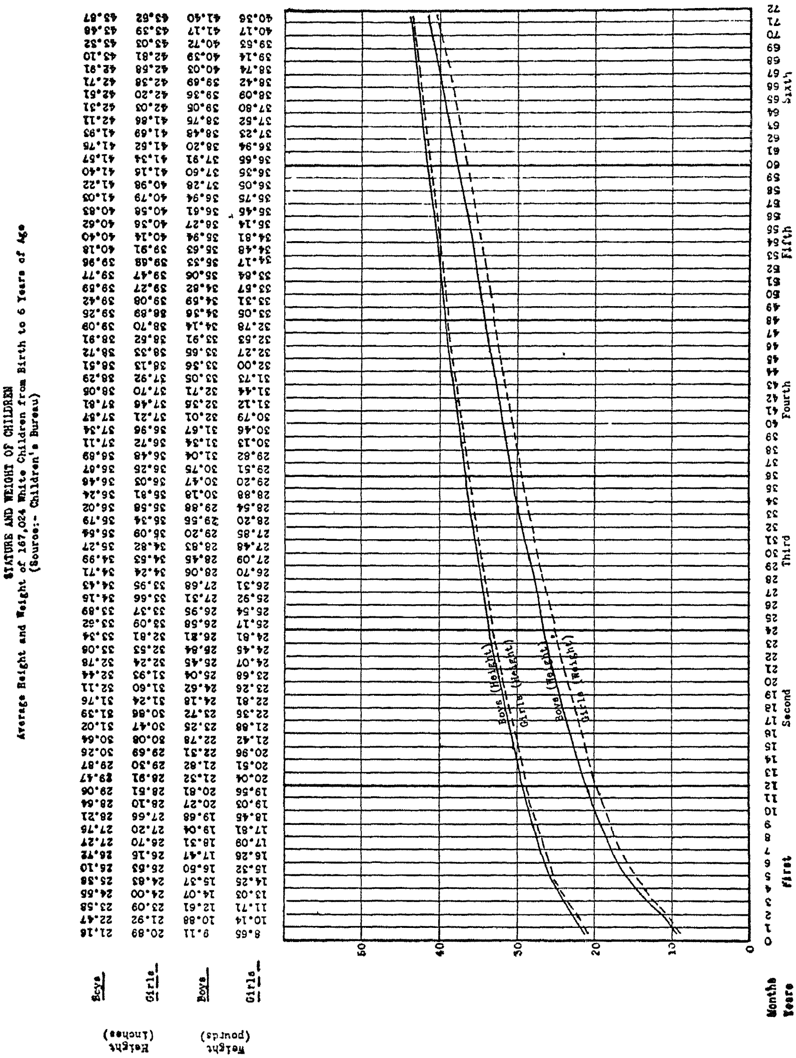


Fig. 295. A Non-cumulable Series.

the connection of the plotted points by a curve is a symbol of the changes of the same phenomena through different conditions, not a short-hand method of indicating different and distinct quantities.

The question of staircase and smoothed curve plotting methods have been given a somewhat lengthy treatment because it has afforded an opportunity to consider the distinctions between discrete and continuous series and period and point data. As a matter of fact if there be but a sufficient number of intervals or groupings in our series, the distinction between staircase and smoothed curve plotting disappears, and both forms of charts become alike and merge into a third,

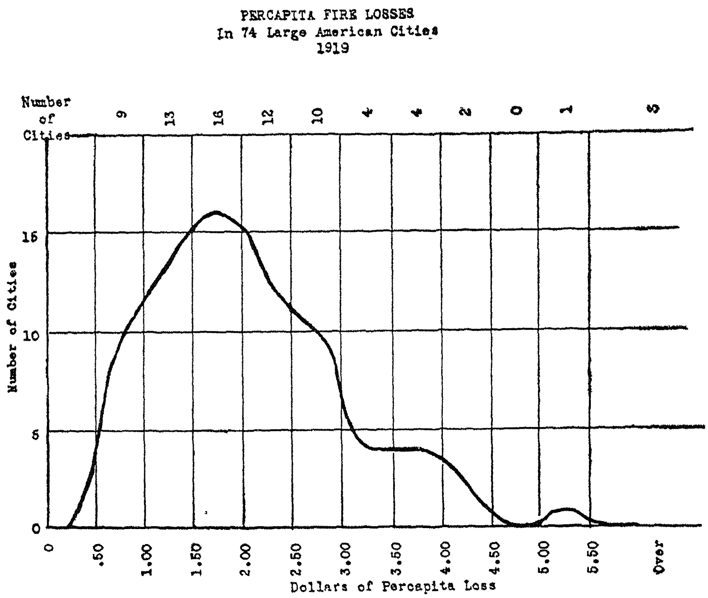


Fig. 296. A Rounded Curve.

the rounded curve. The rounded curve is the true frequency curve and is superior in significance even to the smoothed curve or frequency polygon as the latter is to the staircase (rectilinear or bar-form) curve or histogram. For the rounded curve not only gives gradual change of values between plotted points, but it also gives gradual changes of the rates of change of these values. The derivative of a frequency polygon would be a histogram, but of a rounded curve would be another rounded curve or at least a smoothed one.

We have not, however, laid much emphasis upon the rounded curve, because if the data be sufficiently detailed it will be approximated by either staircase or smoothed curve plotting. Some authorities recommend the artificial rounding

WORKMEN'S COMPENSATION
Delays in Making Payment for Claims in Three States
New York (state and private insurance)
Pennsylvania (state and private, except self insurance)
Massachusetts (private insurance only)
(Source·- Monthly Labor Review)
(Figures show percentage of total cases, 137 in N.Y., 4,093 in Pa. and 186 in Mass.)

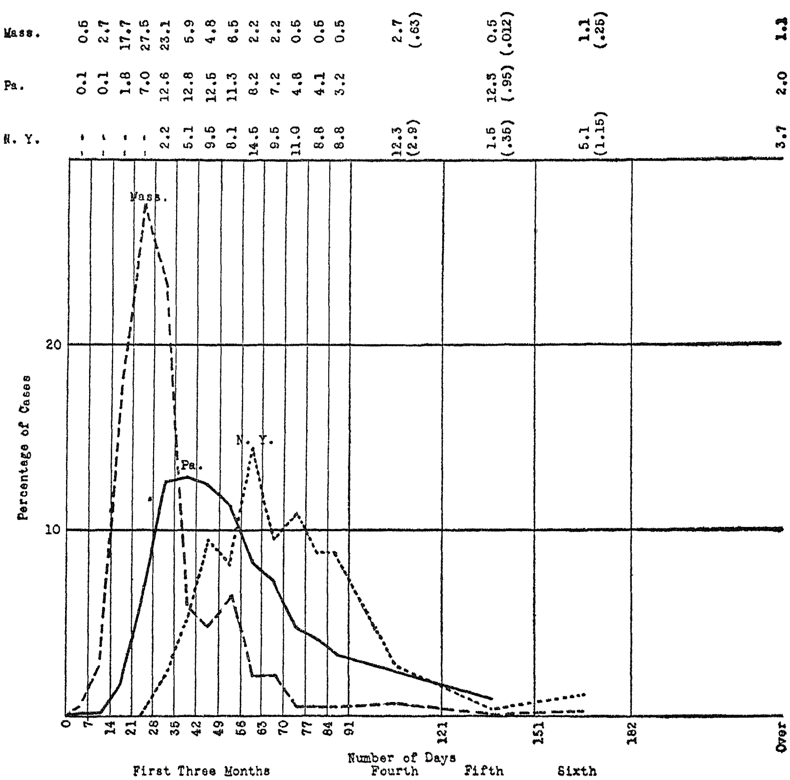


Fig. 297.
Computed averages must be used for the irregular intervals.

of curves.[2] This is to be done either free hand, or by a curving ruler (called by draughtsmen a "French curve"), taking care in either case to pass the curve through all known values (i.e. plotted points), but making the remainder of the curve as little angular as possible. The result is almost always more interesting to the casual reader, obviously because of the more faithful portrayal of the nature of changes from interval to

[2] "The object of smoothing is to eliminate accidental variations and establish normal tendencies."—King, *Elements of Statistical Method*, p. 108.

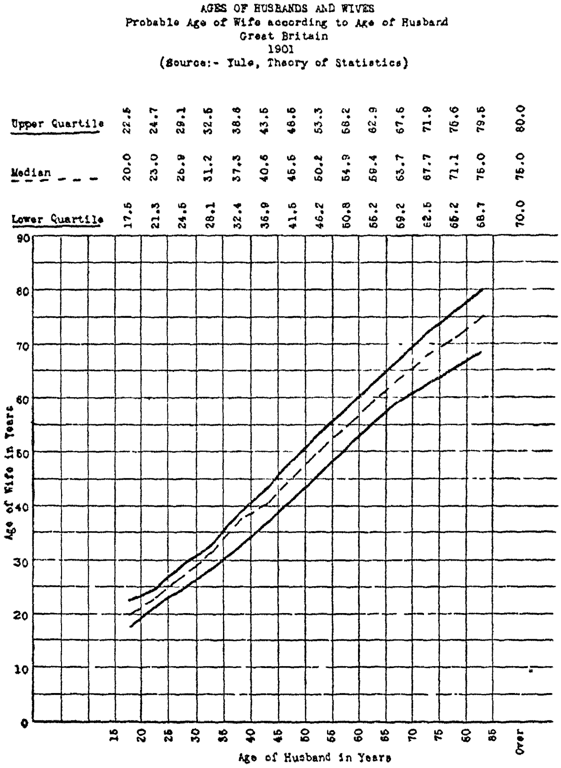


Fig. 298. A Zoned Frequency Curve.

interval. But it is a dangerous practise for the beginner to round his curves artificially, being a wholly inspired embellishment and often leading to slipshod execution. In the early work of the student, the smoothed curve is all that should be attempted, for it is all that the data establishes. And in the research office, it is for the same reason about all that is necessary or safe. In fact, in the research office, it is often sufficient to plot the points only and omit altogether their connections which form the curve, since it is often desirable to superimpose thereon rounded curves of a theoretical nature.[3]

In general, the problems in the graphic presentation of curves arise, first, in the selection of the independent variable

[3] Needless to say, composite curves may be drawn for frequency series as for historical series, in the various forms which have been described (see Chapter XIX). Thus, we may have relative and absolute band-charts, gun-shot plotting, and even vertical and horizontal bar-charts.

FEMALE ACCIDENT MORTALITY RATES
Death-rates per 100,000 of Population of Females of Each Age
For Specified Accidents
United States
1910-1912
(Source - Mortality Statistics, United States Census)

All Accidents	75.2	25.7	11.6	14.6	15.8	15.3	16.5	20.3	22.7	26.6	35.0	45.7	63.6	280.0
Miscellaneous	65.2	21.0	7.6	9.8	11.0	11.2	11.3	14.4	15.5	18.4	23.0	27.7	36.4	87.9
Railways	0.9	0.9	0.9	1.2	1.4	1.4	1.5	1.7	2.1	2.2	2.9	3.3	4.1	6.4
Drowning	3.7	1.6	1.9	2.6	2.2	1.2	1.1	1.1	1.4	1.4	1.0	1.1	1.0	1.8
Falls	6.4	2.2	1.2	1.0	1.2	1.5	2.0	3.1	3.7	4.6	8.1	13.6	22.0	184.1

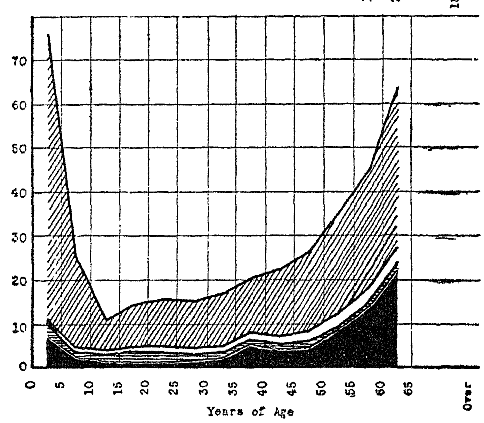


Fig. 299. A Frequency Band-chart.

and the processes of compiling the frequency series; second in the establishment of group limits for the groups in the series and the conversion of irregular intervals into corresponding equivalents; third, in the plotting of data on the ordinates or between them and lastly, in the use of the staircase curve (histogram) or smoothed curve (frequency polygon). To the solution of these problems the distinctions between discrete and continuous data (the integral and graduated variates) and between period and point data bring some assistance, but no set rules of thumb can be given, to which exceptions may not be found. In the wide field of frequency curves, to which the historical curve stands in the relation of a small but important part, the chart-maker and the statistician must rely largely upon native judgment and the precepts of his individual experience.

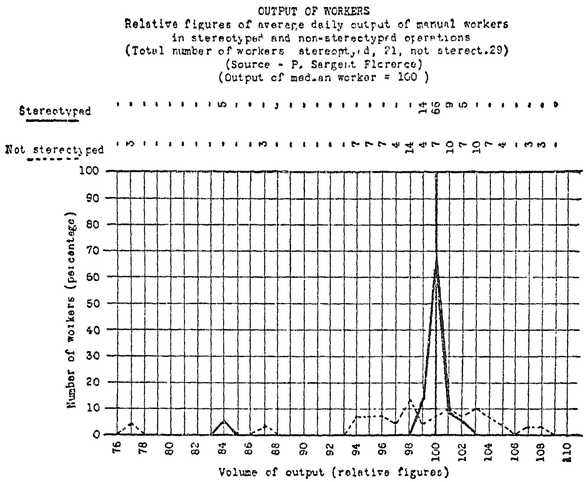


Fig. 300.

A "Relative" Frequency Curve is as possible as a relative historical one

Chapter XXIX

OGIVES

If moving totals are your strongest weapons in the analysis of historical statements, cumulation is your trump card in the analysis of a frequency series. Where the series is continuous (as explained in the previous chapter) the process of cumulation does away with the need for the staircase curve, and gives us a smoothed curve which is far more convenient. And for both discrete and continuous series, the curve of the cumulated data is one which can be easily compared with similar curves, regardless of differences of scale figures or group units. The curve of cumulated frequency is called an "ogive," from its resemblance to the outline of a shoulder. It runs diagonally across the chart, generally in an "S"-shape.

SIZE OF FARMS
United States
1920
(Source:- Census)

Simple series		"Less-than" cumulation	
Acreage	Number	Acreage	Number
Less than 3	20,350	Less than 3	20,350
3, less than 10	268,422	" " 10	288,772
10, " " 20	507,762	" " 20	796,534
20, " " 50	1,503,734	" " 50	2,300,268
50, " " 100	1,474,753	" " 100	3,775,021
100, " " 175	1,449,659	" " 175	5,224,680
175, " " 260	530,795	" " 260	5,755,475
260, " " 500	475,692	" " 500	6,231,167
500, " " 1000	149,812	" " 1000	6,380,979
1000 and over	67,387	Total	6,448,366

Fig. 301. The "Less-Than" Cumulation.

While it would be meaningless to cumulate a historical series backward, and we therefore cumulate historical series only in one direction, it is possible to cumulate a frequency series from either end of the series. If the cumulation begins at the lower end of the data, it is called a "less-than" cumulative, for the sub-total or progressive cumulative figure represents the number of items having less than the maximum qualification of the last added group. If the cumulation begins at the upper end of the series it is called a "more-than" cumulative, the sub-total or progressive cumulative figure representing

SIZE OF FARMS
United States
1920
(Source:- Census)

Simple series		"More-than" cumulation	
Acreage	Number	Acreage	Number
1000 and over	67,387	1000 and over	67,387
500 and less than 1000	149,812	500 " "	217,199
260 " " " 500	475,692	260 " "	692,891
175 " " " 260	530,795	175 " "	1,223,686
100 " " " 175	1,449,659	100 " "	2,673,345
50 " " " 100	1,474,753	50 " "	4,148,098
20 " " " 50	1,503,734	20 " "	5,651,832
10 " " " 20	507,762	10 " "	6,159,594
3 " " " 10	268,422	3 " "	6,428,016
Less than 3	20,350	Total	6,448,366

Fig. 302. The "More-Than" Cumulation.

the number of item having more than the minimum qualification of the last added group. It is often useful to prepare both the "more-than" and "less-than" cumulatives for a frequency series and to plot them both on the chart as well as the series itself.

It is one of the great advantages of the ogive that by its means a frequency series may be graphically presented and analysed whether or not the groups of the series be uniform in size (group-range). There is no labor of calculating values for equivalent groups. All question of staircase curves likewise disappears, even discrete series, when cumulated, being properly shown smoothed. It is another advantage that several

ogives can be easily shown together and compared upon the same chart. The various series need not have uniform and identical group intervals. The ogive is therefore the only feasible method of comparing frequency series which do not have the same groupings. A further benefit is that the many ogives will generally be found to intersect very little, so that the confusion which attends superimposed frequency curves is avoided by ogives.

For analytical purposes it may not be amiss to note that the median of a series is shown by the intersection of the two

EXPECTANCY OF LIFE
For Adults without Tuberculosis
Registration Area, United States.
1910
(Source:- L. I. Dublin: Cost of Tuberculosis)

Years of Age	Average After Lifetime
20	46.6
25	42.3
30	38.1
35	33.9
40	29.9
45	26.0
50	22.1
55	18.5
60	15.2
65	12.1
70	9.5
75	7.2
80	5.4
85	4.0
90	2.9
95	1.9

Fig. 303. An Example of a Frequency Series (So-called) Which Cannot be Cumulated.

ogives, the more-than and the less-than, for the series, or by the value of the abscissae at the intersection of the curve with the ordinate of half the height of the 100% ordinate; while the mode is shown by the portion of the ogive in which the slope is steepest. These are statistical rather than charting conceptions. The median may be described as the middle or central

observation, and the mode as the most common observation.[1] We may also note that two minor variations of the two cumulatives obtain, which depend in part upon the plotting

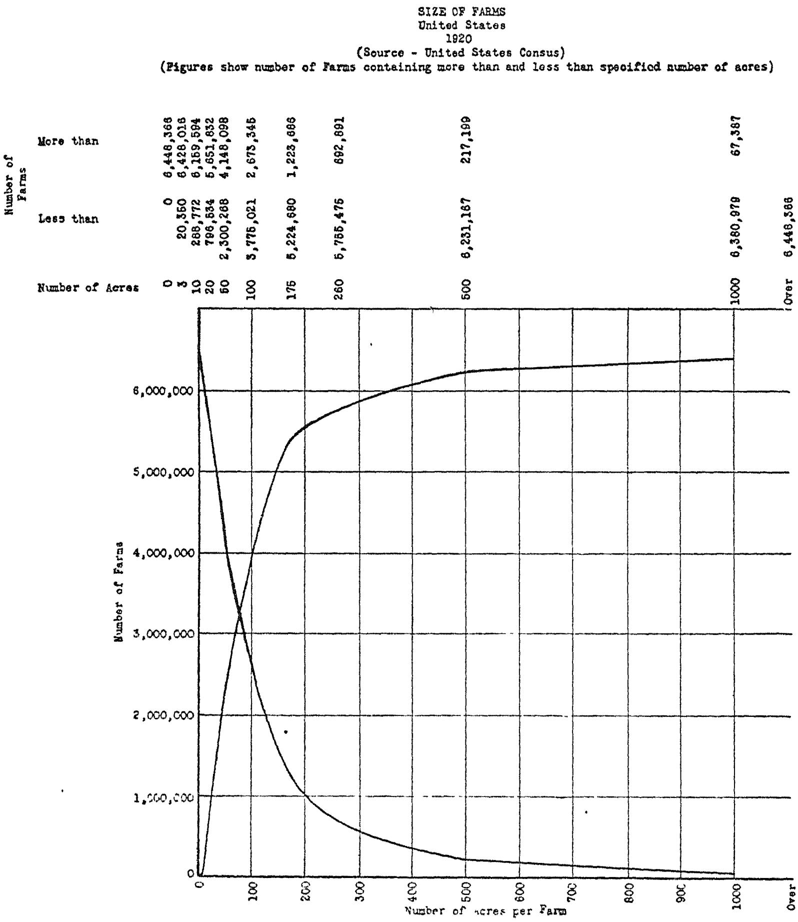


Fig. 304. Two Ogives Are Always Possible.

and in part upon the nature of the data. These variations are, for the "less-than" cumulative, a "less than and including"

[1] Readings from the curve, for the median, decils, quartiles, or percentiles, when secured by interpolation from the curve and not from plotted points on the curve give values, but of course do not give cases. The median case, for example, can be found only by reference to the original data, and exists only if the total number of frequencies be odd. The median value, however, is the intersection of the curve with the 50 per cent abscissa, and is obtained with increasing accuracy as the frequency groups are taken smaller and smaller, and the curve itself plotted in greater detail.

SIZE OF FAMILIES
Number of Children of 1,000 Women
(married at least 15 years and having at least one child each)
British Peerage Statistics
(Source.- Yule, Theory of Statistics)

Corrected Cumulatives	More	973.5	918.5	840	724	588	456	337.5	236	151	87	48.5	25	9	3	1.5	0.5
	Less	26.5	81.5	160	276	412	544	662.5	765	849	913	951.5	975	991	997	998.5	999.5
Uncorrected Cumulatives	"- - Or More"	1000	947	890	790	658	518	394	281	189	113	61	36	14	4	2	1
	"More than - -"	947	890	790	658	518	394	281	189	113	61	36	14	4	2	1	0
	"- - Or Less"	53	110	210	342	482	606	719	811	887	939	964	986	996	998	999	1000
	"Less than - -"	0	53	110	210	342	482	606	719	811	887	939	964	986	996	998	999

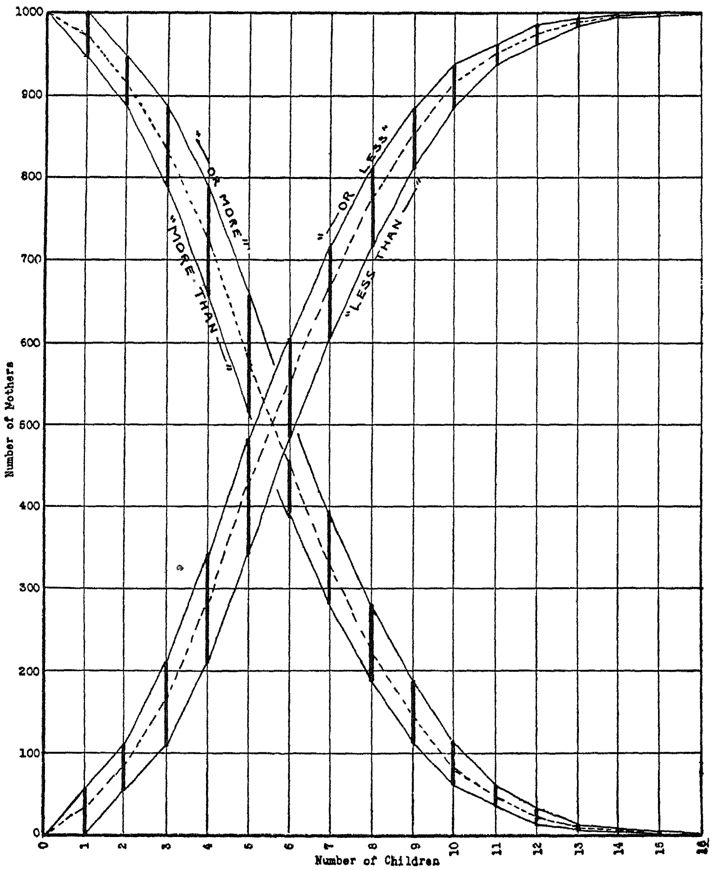


Fig. 305. Showing the Four Possible Cumulations For Point Data.

To find the median or quartiles, etc., graphically, it is necessary to use the dotted line showing the averages of these cumulatives (*i.e.*, the "corrected cumulatives").

cumulative (stated concisely as "–and less") and, for the "more-than" cumulative, a "more than and including" cumulative (stated concisely as "–and more"). They are essen-

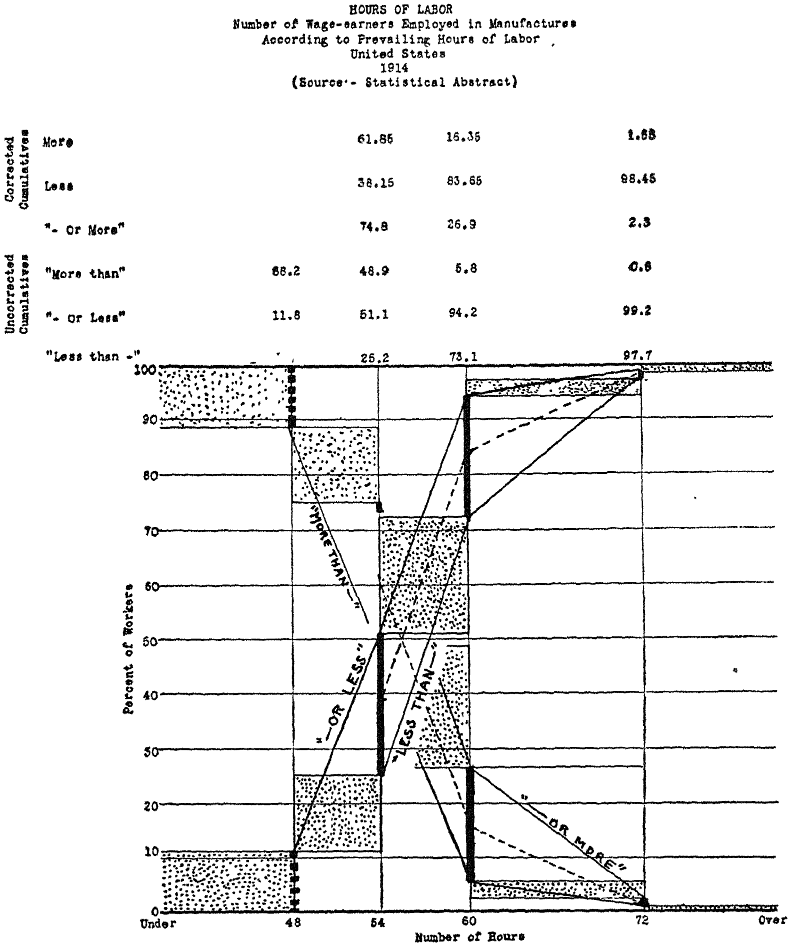


Fig. 306. The Four Possible Cumulations for (Point-and-) Period Data. To find the median or quartiles, etc., graphically, it is necessary to use the dotted lines of the averages between these cumulations (*i.e.*, the "corrected cumulatives").

tially due to differences in the plotting points of data, and are sometimes more suitable for cumulations of discrete data.

It has not generally been observed that the ogive is really simpler in its nature than the frequency curve. Because we have secured the frequency series by a very careful grouping

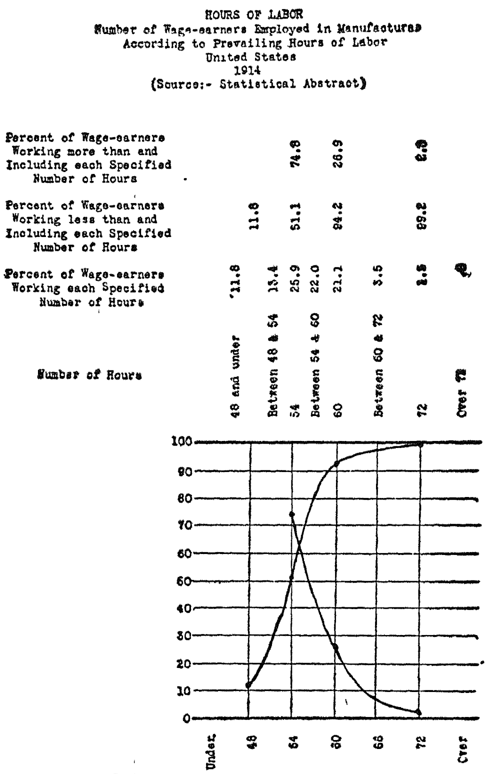


Fig. 307. The Rounded Ogives.

Showing that median and quartiles, etc., cannot be easily found from two uncorrected ogives.

of our original data, arranged in order of magnitude, and have then derived the ogive data from the frequency series by cumulation, we are prone to think of the ogive data as a somewhat more advanced and possibly more puzzling form of statistical series. The fact is, however, that the cumulation merely brings about a reversion to the original data in order of magnitudes, somewhat condensed as a result of the groupings. If we lay off the original data in the form of a bar-chart we will see at once that the ogive is merely a smoothed curve passing through the ends of the bars. It is for this reason that the size of groups or their uniformity is of no importance in making the ogive-chart, smaller groups merely defining the ogive-curve more precisely.[2]

[2] *Cf.* Robert E. Chaddock, in the *American Statistical Association Quarterly*, June 1921, p. 769 ff.

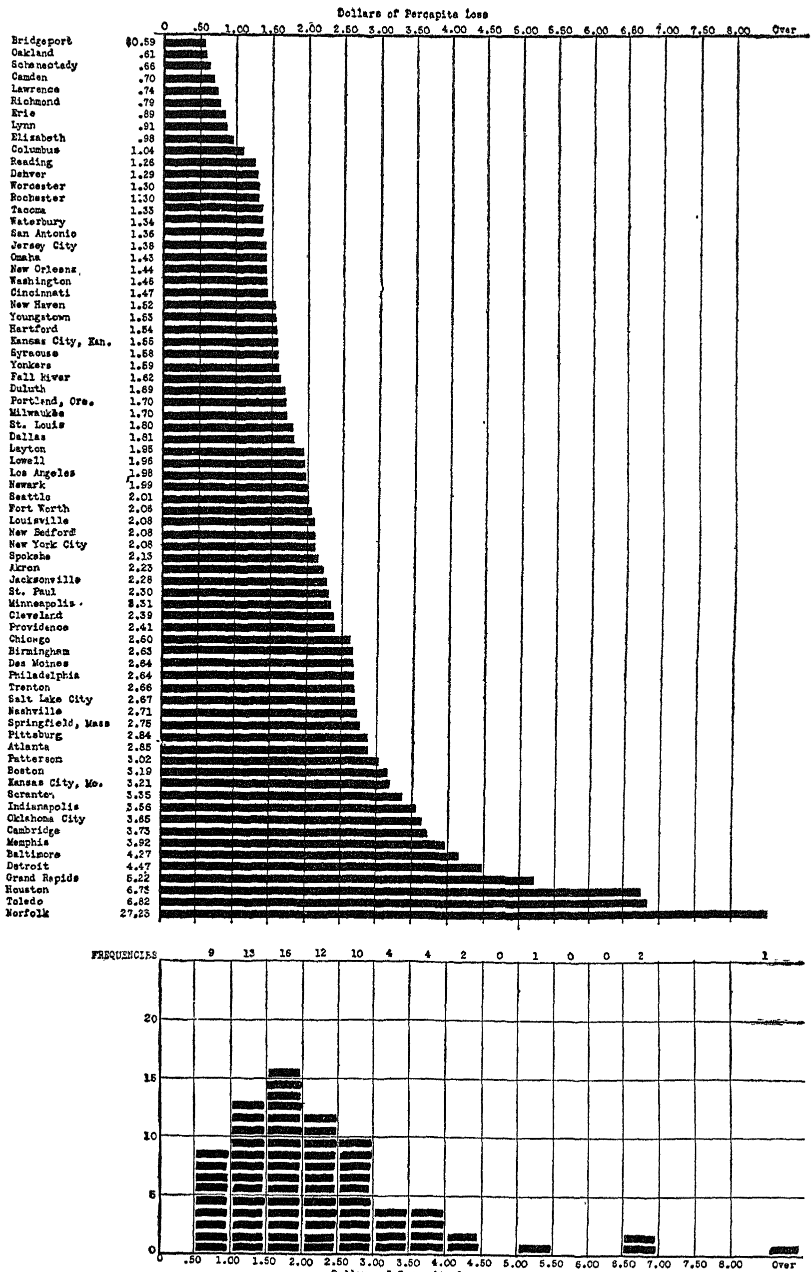


Fig. 308. Evolution of the Ogive (Staircased).

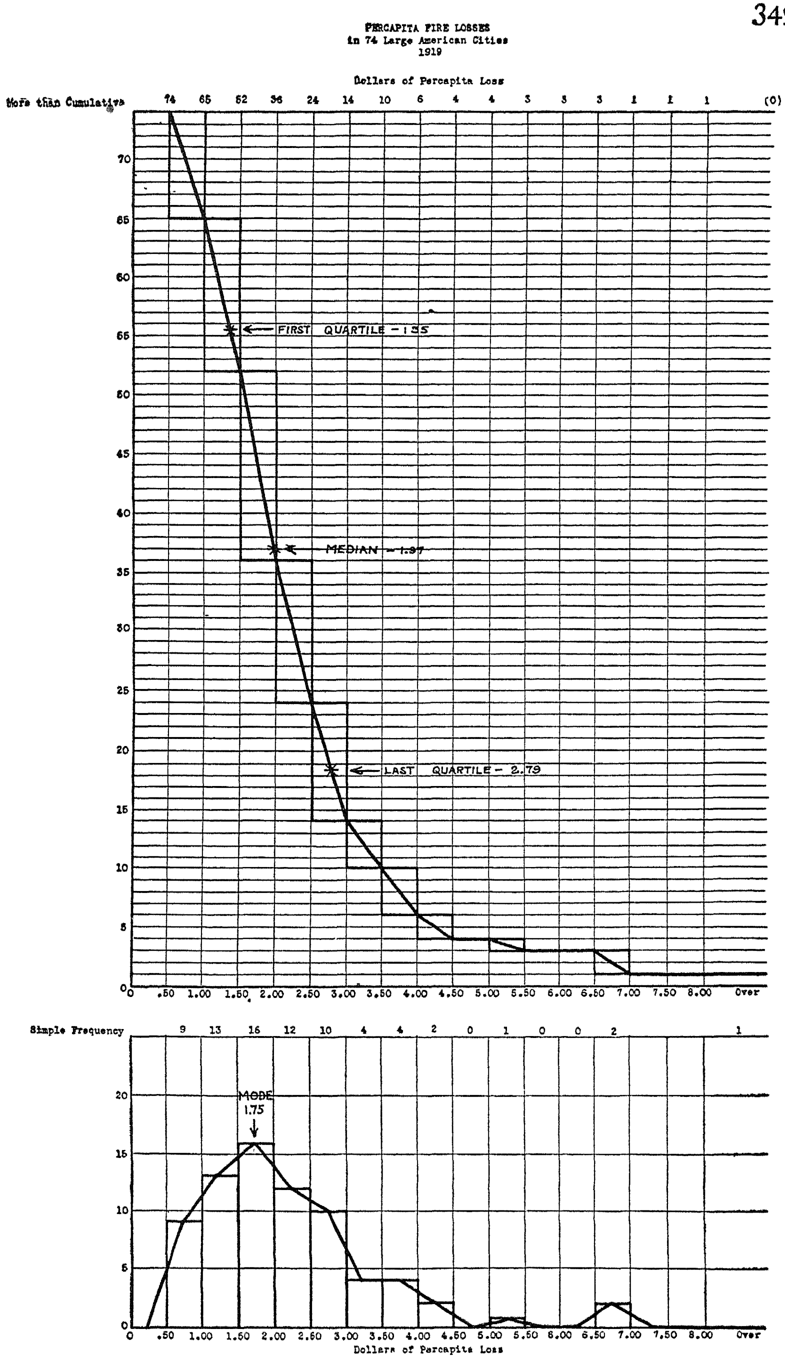


Fig. 309. Evolution of the Ogive (Smoothed).

SIZE OF FAMILIES
Number of Children of 1,000 Women
(married at least 15 years and having at least one child each)
British Peerage Statistics
(Source:- Yule, Theory of Statistics)

Mothers having each specified Number of Children	63	57	100	132	140	124	113	92	75	52	25	22	10	2	1	1
Mothers having up to and including each specified Number of Children	63	110	210	342	482	605	719	811	887	939	964	986	996	998	999	1000
Mothers having more than and including each specified Number of Children	1000	947	890	790	658	518	394	281	189	113	61	36	14	4	2	1

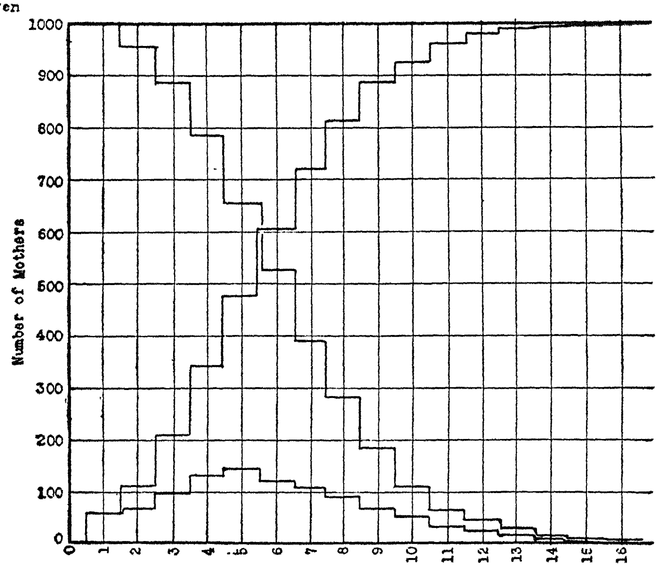


Fig. 310. The Simple Curve and Its Two Ogives (Staircased).

Where individual series are to be analyzed by themselves the horizontal scale for the ogive can be the same as the horizontal scale for the ordinary frequency curve from which the ogive has been derived and for which it has been substituted. The vertical scale, however, will have to be condensed so as to include the total of the entire series. But because a frequency chart is usually designed to show the comparative behavior of the phenomena studied, it is often useful to turn the actual data into percentages and to use on the chart a vertical scale calibrated in percentages. The percentage values are more useful for generalization and ready comparison with other

DURATION OF EMPLOYMENT
Percent Distribution of 1,306 Male and 144 Female Employees on the Payroll
and 2,518 Male and 63 Female Separations who had Served more than Specified Periods of Time
California Sugar Refinery
Active - May 31, 1918
Separated - April 1, 1917 - May 31, 1918
(Source - Paul F. Brissenden)

Active In Service	Males	100	91	87	77	59	46	29	18	14	8
	Females	100	92	86	71	46	18	7	6	4	3
Separated No longer Employed	Males	100	78	65	46	21	12	6	3	2	1
	Females	100	83	70	67	35	22	11	10	10	2
Period of Time		0	1 Week	2 Weeks	1 Month	3 Months	6 Months	1 Year	2 Years	3 Years	5 Years

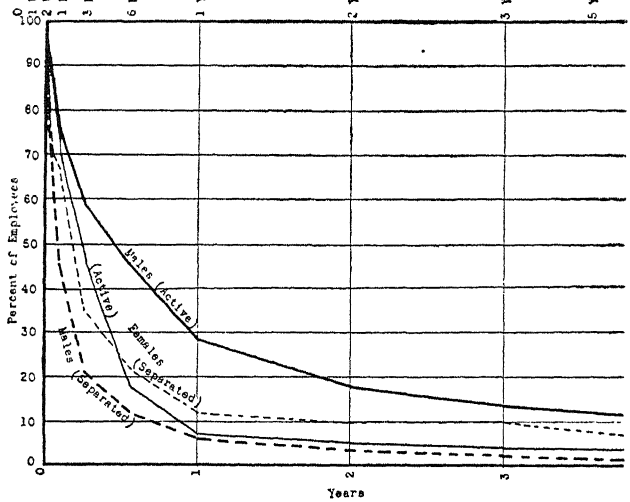


Fig. 311. Relative Data.

ogives, than the absolute scale figures of one particular series of observation or samples. For similar reasons you may find it desirable to turn these group-divisions which form the horizontal scale also into percentages, both in the data and on the chart. In both cases the percentages are percentages of the total or maximum limits of the series. When several frequency series are being compared, and the series differ both in the total number of observations or items in the series and in the group-divisions or group units into which the series is divided, this little trick of turning all readings into percentages may be very useful, as by its means you can chart the ogives or cumulatives of all the series upon uniform chart-fields. The fields on which the curves are to be plotted should generally be

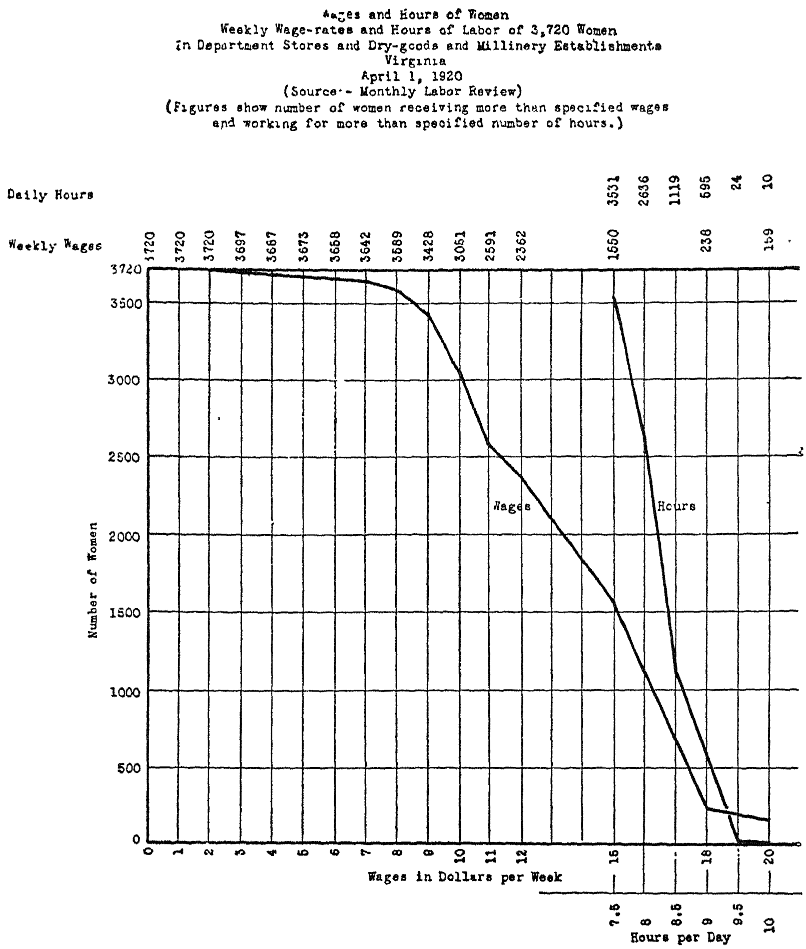


Fig. 312. Comparison of Absolute Data is Sometimes Difficult.

square, running from zero to one hundred per cent along both axes of the chart. Needless to say, the fields should be uniformly positioned upon the sheets of paper so that the various charts to be compared may be freely subjected to "light analysis," that is, to the method of analysis which consists of holding two or more charts together up to the light to detect the variations of their curves.

In the ogive chart we first meet with a type of chart which illustrates at the same time two different sets of figures for the same curve. There should therefore be space for data

not only above the chart but to the right of the chart, and the chart field should not be placed close to or near to the right-hand margin of the paper as was the case in historical curves. The data above the chart is obviously the original data of the cumulative from which it is plotted, each data figure being placed above the ordinate or corresponding scale figure on the

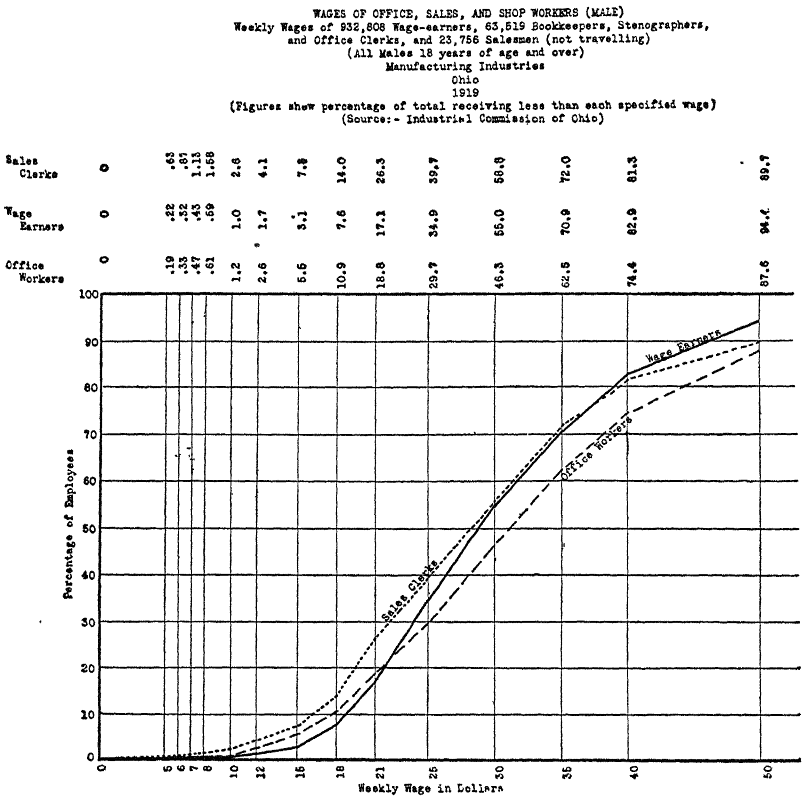


Fig. 313. Comparison of Relative Data is Easy.

x-axis of the chart. The data to the right of the chart field will be secondary or derived data, obtained by taking the readings or values of the curve at each of its intersections with the abscissae or horizontal rulings, using the corresponding scale figures of the y-axis for the new stubs, and taking the corresponding values along the x-axis as the new or derived data. This secondary data forms a new table of the same phenomenon rearranged so that the second variable or de-

pendent has become in a sense the independent one and its values would appear as the stubs in a retabulation.

The ogive-chart is excellently adapted for the process of interpolation. The derived data just described are an example of this use of the chart. Interpolation, of course, is the name given to the process of reading new values between originally given values. Thus, by means of the ogive-chart, originally incomplete data can be filled out with interpolated figures. But the interpolated figures of course do not have the same

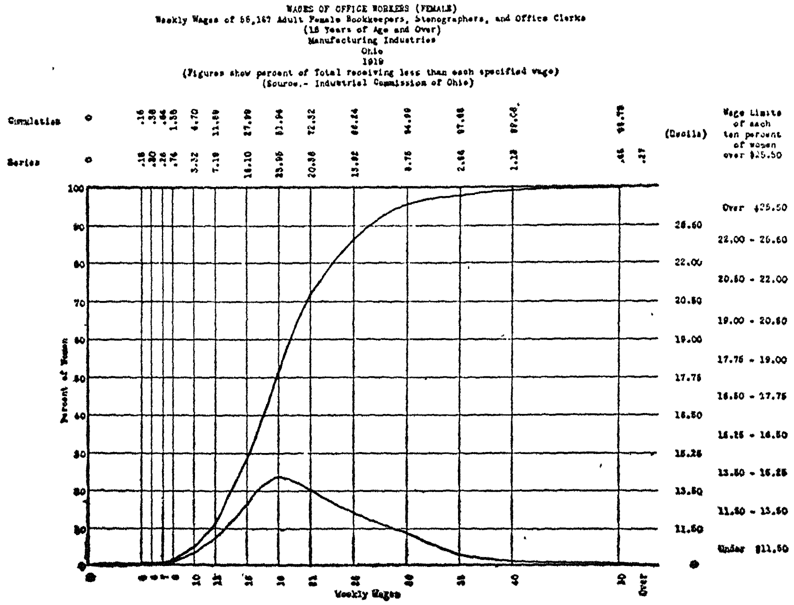


Fig. 314. Secondary Data at the Right.

degree of accuracy as the original data, being made on the theory that the line drawn between the plotted points of the original data has been correctly drawn. In spite of their possible inaccuracies, they are often extremely useful in reducing otherwise incomparable frequency series to comparable group units or to uniform percentages. Where the chart itself is to be used, the interpolation is not necessary, the connecting lines which form the curve being in themselves plottings of interpolated values. But when the chart will not be presented in the final report or summary of the case, the interpolated figures are necessary, the interpolated figures being obtained from the chart before the chart itself is discarded.

The ogive-chart is one of the most important and generally useful of the non-historical chart-forms and we will have occasion to return to it in the future with various elaborations and improvements.[3]

[3] It is thoroughly regrettable that statistical practice has so consistently considered the range of a frequency series to be the independent variable and its frequencies the dependent. For while this is logical enough in the simple curve, it introduces into the ogive or cumulated curve, absurdities which make the latter not only unnecessarily obscure to the layman, but also brings about an unjustifiable violation of the primary rule in curve charting that the plot of all points should be independent x, dependent y.

In the simple frequency curve, we are interested only in frequencies and they are obviously and properly dependent. The reduction *ad initium* of this chart to a bar-chart would, as has been seen, require the use of vertical bars, and the curve is but the short-hand connection of the tops of these bars.

A very different case is that of the cumulated series. Here the frequencies are only in an immediate sense dependent; in the last analysis, they are independent and the range is dependent. As has been seen, the return of this curve to a bar-chart would require the use of horizontal bars—a fact which clearly illustrates the really dependent nature of the range.

Statisticians have, however, so short-sightedly adopted the ogive as a derived chart (by cumulation) from the simple frequency curve, that they have followed its arrangement of the variables on the chart; and as the ogive has been confined to use by statisticians, the practice has become so settled that to advocate a change at this time would seem only to be adding confusion to a science which, more than all else, is in need of standardization.

Life does not, however, always confine itself within academic rules, the roundabout path must always be supplemented with a fence; and we venture the prediction that in time, as the ogive comes into commercial use, this arrangement will be scrapped and the ogive plotted on x-frequencies, y-range, passing first through a period of confusion which we do not seek to bring about. But he who must tell his story clearly or not at all will drop the old arrangement. A step in this direction, though perhaps un'concious, is that of the publishers of probabilities paper for ogives, who probably only by accident or for convenience have calibrated their scales so as to give x-frequencies, y-range. The chart in this form is more intelligible to the layman.

CHAPTER XXX

LORENZ CURVES

Neither the frequency curve nor its ogive have that peculiar tang of popularity, that engaging frankness which appeals to the "average man." If we consider the ogive the simpler of the two, since it is merely a curve passing through the ends of horizontal bars, then indeed it is a curiously unscientific chart, in which the usual position of the dependent and independent variables is reversed. The frequency curve is then a short-hand method of arraying these, with a large degree of chance in its formation when various groupings have different results. And if we consider the ogive as a cumulated frequency curve it then involves all the obscurities of the simple frequency curve, augmented by further complexities of its own. We

OUTPUT OF FACTORIES

Number of Manufacturing Establishments
of specified sizes
(size being measured by value of products)
United States
1914
Source:--U.S.Census

Specified Value of Products per Establishment	Number of all Establishments having specified value of products	
	Number	Percent
Less than $5,000	97,061	35.2
$5,000 - $20,000	87,931	31.9
$20,000 - $100,000	56,814	20.6
$100,000-$1,000,000	30,166	10.9
$1,000,000 and over	3,819	1.4
TOTAL	275,791	100.0

Fig. 315. The First Measure—By Count of Items.

now take up, therefore, a curve which has more popular elements in it, with a consequent sacrifice of statistical detail.

OUTPUT OF FACTORIES

Value of Products of Manufacturing Establishments
of specified sizes
(size being measured by value of products)
United States
1914
Source: -- U.S.Census

Specified Value of Products per Establishment	Value of Products of all Establishments having specified value of products	
	Dollars	Percent
Less than $5,000	233,381,081	1.0
$5,000 - $20,000	905,693,168	3.7
$20,000 - $100,000	2,550,229,411	10.5
$100,000-$1,000,000	8,763,070,135	36.1
$1,000,000 and over	11,794,060,929	48.6
TOTAL	24,246,434,724	100,0

Fig. 316. The Second Measure—By Count of Units.

All frequency distributions afford two possible series for precisely the same data. The first and more usual series is the count of items in each group of the distribution, the second

OUTPUT OF FACTORIES

Number and Value-of-Products of Manufacturing Establishments
of specified sizes
(size being measured by value of products)
United States
1914
Source:- U.S.Census

Specified Value-of-Products per Establishment	All establishments having specified value-of-products			
	Number		Value-of-products	
	Number	Percent	Dollars	Percent
Less than $5,000	97,061	35.2	233,381,081	1.0
$5,000 - $20,000	87,931	31.9	905,693,168	3.7
$20,000 - $100,000	56,814	20.6	2,550,229,411	10.5
$100,000-$1,000,000	30,166	10.9	8,763,070,135	36.1
$1,000,000 and over	3,819	1.4	11,794,060,929	48.6
TOTAL	275,791	100.0	24,246,434,724	100.0

Fig. 317. Both Measures.

and alternative series is the count of the units of measurement attributed to these items. Thus a classification of farms by their size (in acres of land) can show us either the number of farms of each size or the aggregate number of acres in the farms of each size. The data of cities classed by their population may count either the number of cities of each specified number of inhabitants, or it may count the inhabitants residing in these cities. The Census of the United States, in its analysis of the manufacturing establishments of the country, according

OUTPUT OF FACTORIES

Number and Value-of-Products of Manufacturing Establishments
of specified size
(size being measured by value-of-products)
United States
1914
(Note:-- All data in percentages of total or aggregate)
Source:-- U. S. Census

Specified Value-of-Products per Establishment	All Establishments having specified value-of-products	
	Number	Value-of-Products
Less than $5,000	36.2	1.0
$5,000 - $20,000	31.9	3.7
Less than $20,000	67.1	4.7
$20,000 - $100,000	20.6	10.5
Less than $100,000	87.7	15.2
$100,000-$1,000,000	10.9	36.1
Less than $1,000,000	98.6	51.3
$1,000,000 and over	1.4	48.6
Any value whatever	100.0	100.0

Fig. 318. Cumulating the Percentages.

to their employees, gives both the number of establishments and the aggregate number of employees in such establishments; in its analysis by value of products, gives both the number of establishments and their aggregate value of products. Examples might be multiplied without end, for whenever we distribute the items of any phenomenon into groups upon the basis of some units for measurement, we are then at liberty to count either the items themselves or their units of measurement, group by group.

The thought, therefore, occurs to us that a chart could be made in which one of these series serves as the independent variable for the other, and in which the two values for each

OUTPUT OF FACTORIES

Value-of-Products and Number-of-Establishments
of Manufacturing Establishments of specified size
(size being measured by value-of-products)
United States
1914
(Note:- All data in percentages of total)
Source:- U.S.Census

Specified Size as shown by Value-of-Products per Establishment	Establishments having specified value-of-products	
	Number-of-Establishments	Value-of-Products
Less than $5,000	35.2	1.0
Less than $20,000	67.1	4.7
Less than $100,000	87.7	15.2
Less than $1,000,000	98.6	51.3
Any value whatever	100.0	100.0

Fig. 319. Data For the Lorenz Curve.

group of items are made the co-ordinates of plotted points. Obviously, if we are not to have the curve which connects

OUTPUT OF FACTORIES

The Value of Products of
specified groups of Manufacturing Establishments
United States
1914
In percentage figures
(Note:-- All groups composed of establishments having least value of products.)
Source:-- U S Census

Number of Establishments	Value of Products
Percent	Percent
35.2	1.0
67.1	4.7
87.7	15.2
98.6	51.3
100.0	100.0

Fig. 320. Data to Plot the Lorenz Curve.

these points moving backward and forward as well as up and down, we must cumulate the series which is to be used as the

independent variable, and since it will be found that the curve has greater significance when both are cumulated, we invariably cumulate both series. It is then a matter of indiffer-

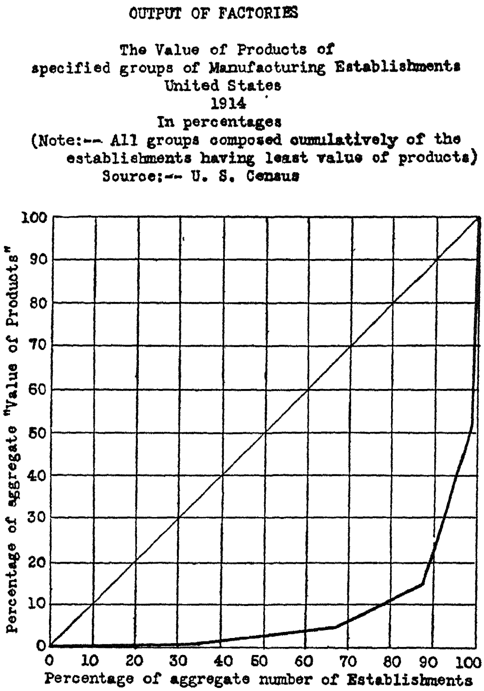


Fig. 321. The Lorenz Curve.

ence which series be used as the independent one, and either series may be plotted upon either axis of the chart. It will be seen that the chart is closely related to the ogive, since it uses

OUTPUT OF FACTORIES
United States
1914.

The	largest	10	percent	of	the	factories	produce	78	percent	of	the	values;	the	smallest	90,	only	22.
"	"	20	"	"	"	"	"	87	"	"	"	" ;	"	"	80,	"	13.
"	"	30	"	"	"	"	"	93	"	"	"	" ;	"	"	70,	"	7.
"	"	40	"	"	"	"	"	96	"	"	"	" ;	"	"	60,	"	4.
"	"	half		"	"	"	"	97	"	"	"	" ;	"	"	,	"	3.
"	"	60	"	"	"	"	"	98	"	"	"	" ;	"	"	40,	"	2.
"	"	70	"	"	"	"	"	99	"	"	"	" ;	"	"	30,	about	1.
"	"	80	"	"	"	"	"	$99\frac{1}{3}$	"	"	"	" ;	"	"	20,	"	$\frac{1}{2}$.
"	"	90	"	"	"	"	"	$99\frac{3}{4}$	"	"	"	" ;	"	"	10,	"	$\frac{1}{4}$.

(Note:--"Large" and "small" refer to the size of the factory as measured by the value of its products.)

Also

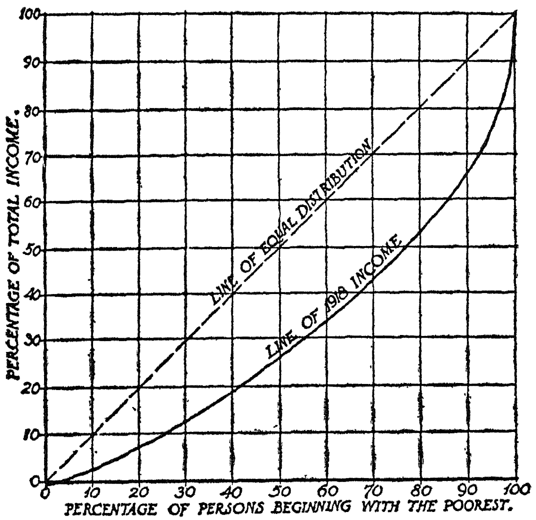
10 percent of the values are produced by the largest $\frac{1}{2}$ percent of factories; 90 by the smallest $99\frac{1}{2}$
20 " " " " " " " " " 1 " " " ; 80 " " " 99.
etc.........

Fig. 322. What the Lorenz Curve Tells the Layman.

a cumulated series; it will also be seen that the chart omits altogether the classes or groups in which the data has been collected. Lastly, to produce a uniformity of these charts, and to facilitate the comparison of different distributions upon the same chart, all items are turned into percentages of the totals, and plotted upon percentage scales along both axes.

When this type of curve is drawn upon a square field, with equal percentage scales upon each axis, it takes the shape of an archer's bow, and the curvature of the bow has a peculiar

LORENZ CURVE SHOWING THE DISTRIBUTION OF INCOMES IN 1918.



From "Income in the United States," by the National Bureau of Economic Research, by permission.

Fig. 323. The Familiar Example.

significance as an index of dispersion in the original distribution. For a little thought will show that a uniform distribution in which all items are alike will yield not a curve, but a straight line. The first ten per cent of the "population" in the series of personal incomes, for example, if all incomes were equal, would have ten per cent of the total income of the country, the first twenty per cent would have twenty per cent of the total income, and so on. Hence, the distance between the curve and the straight-line diagonal indicates the degree in which the series is removed from a perfectly uniform dis-

tribution—a feature which statisticians call dispersion or scatteration.

The Lorenz curve, as this form of chart has come to be known, has not been much used except in the analysis of income and wealth distribution, but it is obvious that it is capable of use for any and all frequency series. It is simple

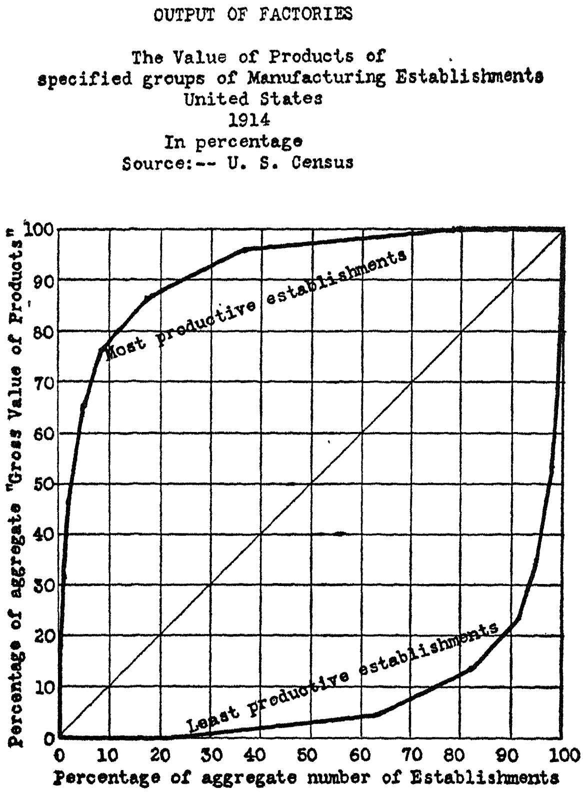


Fig. 324. Two Curves of the Same Data By Using Both "More-than" and "Less-than" Cumulatives.

and popular in its appeal, without being in the least inaccurate or meaningless. It has certain advantages in the emphasis it throws upon dispersion and unequal distributions. Its chief disadvantage is in the omission of the group-by-group data for the series it illustrates, but this data is more in the nature statistical detail, and does not belong to what may be called a summary analysis of a distribution; to the average man such detail is confusing rather than helpful, while the results of the

dispersion, which this chart shows, form in his mind the meat of the matter.[1]

The principles of the Lorenz curve can, however, be extended to innumerable comparisons between frequency series. It is not necessary that the two series compared be the two alternative forms for the same data. Though the latter is usually the sounder practice, there may be occasion to bring

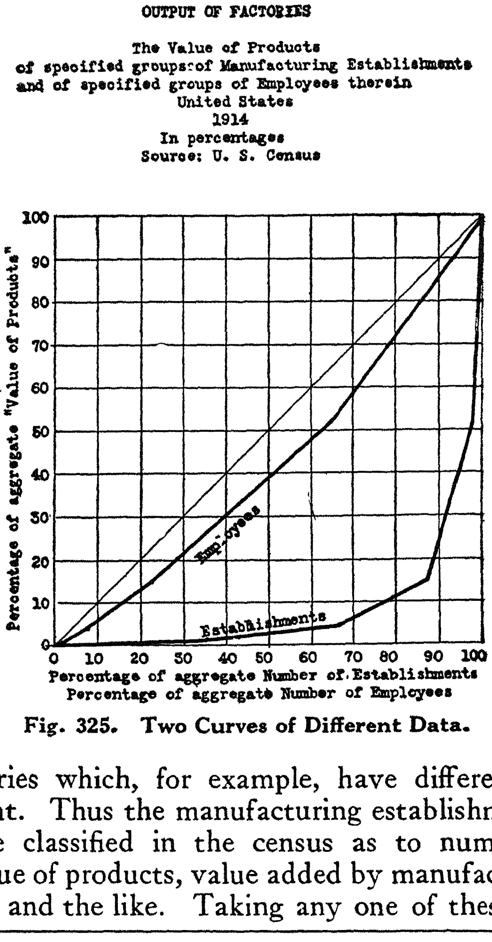


Fig. 325. Two Curves of Different Data.

together series which, for example, have different units of measurement. Thus the manufacturing establishments of the country are classified in the census as to number of employees, value of products, value added by manufacture, horsepower used, and the like. Taking any one of these classifica-

[1] As no data can be easily appended to the Lorenz curve, unless we elect to give readings of the curve at various points (in which case data belongs at both the top and on the right side of the chart, to give readings for both variable scales), it is generally sufficient to append to the Lorenz curve the percentage cumulations from which the curve has been drawn. These afford to the inquisitive full details for the group-by-group distributions which the chart itself does not show.

tions, the census gives the other features just mentioned, for the establishments forming each group in the classification. The true Lorenz curve will then bring together, for example, the number of establishments having a specified value of products and the group value of the products of these establishments. If we like, however, we can bring together the value of products of each group and the number of employees attached thereto, or the value added thereby, or any other

OUTPUT OF FACTORIES

"Value of Products" and "Value Added by Manufacture"
of specified groups of Manufacturing Establishments
and of specified groups of Employees therein
United States
1914
In percentages
Source: U. S. Census

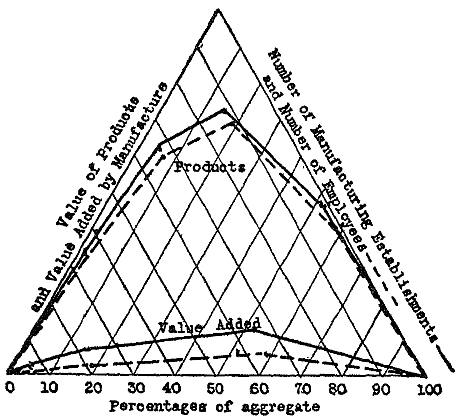


Fig. 326. The Logical Form is Triangular.

feature we desire to show. This forms a pseudo-Lorenz curve which, though slightly more complicated in principle, has the same popular features.

Popularity is the main feature of the Lorenz curve, but it is not without its scientific significance. As already remarked, the deviation of the curve from a straight line shows the dispersion or scatteration of a series. For this reason, it would seem useful to plot the Lorenz curve upon triangular or tri-axial ordinates, by omitting the useless half of the square chart-field, and to make the stright-line diagonal the base of the chart in order to emphasize the deviation of the curve from the straight-line diagonal. A second feature of the

Lorenz curve is that the lack of similarity between the two terminal parts or "tails" of the curve, indicates what statisticians call skewness in the data. In these ways, this form of chart is a useful tool for the technician, and yields an intelligible message of details in which he is interested and which will escape the lay reader. Primarily, however, the Lorenz curve is popular—the one and only way of making an interesting picture of a frequency distribution for the average man and of bringing strongly home to him the practical aspects of the frequency distribution presented.

PART III. RATE-OF-CHANGE ANALYSIS

Chapter XXXI

THE GENEALOGY OF NUMBERS

The first mathematical operation in the world was probably more difficult for its discoverer, than the most complicated mathematical processes are for us today. Prehistoric man was able to master that initial operation and hence you are confronted with the disconcerting question, "are you intellectually weaker than the cave man, the stone-age man and the iron-age man?" If you admit the charge, or if you have already mastered higher mathematics, you should skip this chapter and continue in the straight and narrow road of charting, for in the first case you will not understand it and in the second you will not need it. The chapter is by way of being a comic interlude in which the reader is invited to wander down a by-way into pure mathematics, which will give him a theoretical understanding of the charts which follow. For a practical working knowledge of them, this theoretical understanding is not needed, but it will be a source of abiding satisfaction to him and will incidentally raise his batting average against charting errors and mistakes.

The first mathematical operation was that of counting off or numbering. That is to say, standing in the middle of a road, you walked forward and your first step was your "first" step, your next was your "second" step, your next was your "third" step, and so on. The implements are called ordinals ("first," "second," "third," etc.). The result of this counting off or numbering is to give you the number of steps you have taken, that is, measurement or mensuration. The measurement comes in the form of what is called a cardinal number ("one," "two," "three," and so on). Hence note that you have the following:

Materials:	Distinct items
Operation:	Counting off or numbering
Result:	Ordinal numbers

Materials:	Ordinals
Operation:	Measurement or mensuration
Result:	Cardinal numbers

Several centuries may have passed before another bright young chap came along who was a little less hairy than his ancestors and had a little higher forehead. He discovered that if he walked five miles one day and five the next, it was the same as walking ten miles in all, the interesting thing being that the same two cardinals always made the same third cardinal. This made possible the great and fundamental law upon which our civilization is said to rest, that two and two make four. The operation is known as "addition." It is a sort of multiple measuring. Its result is a "sum." An inverse operation exists which is called "subtraction," the result of which is called a "difference." In this inverse operation we first meet with what are called "negative numbers" giving rise to a conception that numbers can sometimes be either "positive" or "negative." And in this inverse operation we must distinguish the two numbers operated on, calling the first the "minuend" and the second the "subtrahend." Now note that you have the following:

	Direct	*Inverse*
Materials:	Terms	Terms
1st material	———	Minuend
2nd material	Increment	Subtrahend
Operation:	Addition	Subtraction
Result:	Sum	Difference

Again many centuries elapse before the third act. This time the inventor discovers that if he walks five miles every day for five days, it is the same as walking twenty-five miles in all, and that no matter how often he repeats the operation the result will always be the same for each two given numbers. From this we discover the law that two times two make four, and with it the multiplication table. The process is called multiplication and is a sort of multiple adding. Most of our so-called multiplying machines are built on this principle, the operator simply turning the crank of the machine often enough to add the quantity the right number of times. The reverse process is called "division" and is a sort of multiple subtraction. The result of "multiplication" is called a "product," of division a "quotient." And the reverse process, division, gives

rise to a new type of number called a "fraction," showing us that numbers can sometimes be "integrals" or whole numbers and sometimes "fractions" or part numbers. And remembering the negative numbers we met, we also find negative products and quotients or fractions. Note now that we have the following:

	Direct	*Inverse*
Materials:	Factors	Factors
1st material	Multiplicand	Dividend
2nd material	Multiplier	Divisor
Operation:	Multiplication	Division
Result:	Product	Quotient

Again a long time passes. The fourth act begins. Some one says: A five-mile distance walked by five persons, total walking, 25 miles (in which 5 is used twice as a factor), on 5 different days, total 125 miles (in which 5 is used three times as a factor), in five different cities, total 625 miles (in which 5 is used four times as a factor) and in five different countries, total 3125 miles (in which 5 is used 5 times as a factor). Now, says he, instead of writing "$5 \times 5 \times 5 \times 5 \times 5$" why not write 5^5 and be done with it. His invention, you see, is clearly one of notation. His process is called "raising to a power," his result being a "power" of the original number. It is a sort of multiple multiplication. The inverse operation is called "reducing to a root" the result being a root of the original number. It is a sort of multiple division. We use the method when we say that the square (or second power) of two is four. The little number up in the corner is called the exponent. When it is in the righthand corner it signifies raising to a power, and when on the lefthand side in a radical sign it signifies the inverse operation of extracting a root. Again the raising to a fractional power also signifies the inverse process. And in the inverse process we meet with the square root of negative numbers, which we call "surds" or "irrational numbers." And we also find negative exponents. Now note that you have:

	Direct	*Inverse*
Operation:	Involution	Evolution
Materials: 1st	Number	Number
2nd	Exponent	Fractional exponent
Result:	Power	Root

Here, the author, in the role of stage manager, must step out in front of the curtain, with a little speech of apology. The play would progress better had not the playwrights, that is the mathmeticians, been badly put to it to find new names and symbols for their operations. They have progressed bravely up to this point. Thus reviewing their work you find:

Counting off: 1st, 2nd, 3rd, 4th, 5th, etc., gives us Measurement; 1, 2, 3, 4, 5, etc.

Multiple Measuring: 1, 2, 3, 4, 5, and 1, 2, 3, 4, 5, gives us Addition; $5+5=10$.

Multiple Addition: $5+5+5+5+5$ gives us Multiplication; $5\times 5=25$.

Multiple Multiplication: $5\times 5\times 5\times 5\times 5$ gives us an Involution; $5^5=3125$.

So far, they have given us a marvelous system for the easy notation of their ideas. You will notice that the phrase 5^5 is a highly compressed expression, which would otherwise have to be written $5\times 5\times 5$ . . (to 5 times) or $5+5+5+5+5$ . . . (to 625 times) or $1+1+1+1+1$ . . . (to 3125 times). But we warn you that this simplicity is at an end. Examine the expression, $3125=5^5$. Substitute for it the general algebraic expression $A=B^c$. This describes A as the *C-th* power of B. From it you can readily derive the expression for B, as follows: $B=\sqrt[c]{A}$, that is B is the *C-th* root of A. But they have no convenient symbol for C, the exponent of the power to which B must be raised to equal A, and they can only give you a cumbersome word by which you can describe C as—but wait and see.

The fifth act, with which, so far as we are concerned, the play should end happily, opens with a young man who discovers that the fifth power of five is the same thing as multiplying together the second and third powers of five, and that, in general, to multiply two powers of the same number together you need merely add their exponents, thus $B^c\times B^d=B^{(c+d)}$. Likewise, to divide a power by another power of the same number, you need merely subtract their exponents, thus $B^c\div B^d=B^{(c-d)}$. Whereupon he promptly says, let us change all numbers in the world into powers of one common and universal base number and then we shall be able to substitute for the lengthy tedious process of multiplication and division, the simple and easy process of addition and subtraction. Instead of multiplying

together a long series of large numbers we would only need to add their corresponding exponents of this universal base. The discovery was in the nature of a miracle. To add exponents instead of multiplying powers! The process has been accounted one of the nine wonders of the world. And it's as easy as falling off a log!

Then this excellent discoverer has to spoil his work by using a long and terrifying name for his process. For he calls it *logarithmation*. He calls his exponents of the common base, *logarithms to that base*. He calls his table of exponents a table of logarithms. He cannot think up a new symbol and when asked what c is in the equation, $A = B^c$, he writes:

$$c = log_B A$$

(This is read, "*c* equals *log A to base B*".) For he abbreviates his long word logarithm by the short word *log*. But it is no use. The public has decided that the use of logarithms is not as easy as falling off a log. All of which goes to show that there is something in a name after all. The public fell off the logs long ago and has been off them ever since. After this unhappy *denouement*, we introduce the following pageant, as additional entertainment to an audience which has sat faithfully through five tedious educational acts.[1]

Marshal before your eyes the countless myriads of numbers known to man. (For the sake of simplicity consider the whole numbers only, forgetting for the moment the fractions.) Arranged in single file from zero out into infinity, it would take forever for their procession to pass. For there is literally no end to them. Marching by at the rate of one at every tick of the clock, the first ten thousand would pass in an hour. And marching day and night, after four days one million would appear. But it would be ten years before the first number of ten digits, the first billion number, comes into view. And as to the trillion, that would not yet have appeared if the parade had begun before the pyramids were built. Yet the trillion is no longer a stranger to financial circles and is a poor small thing in the world of science.

Now hovering over the shoulder of every one of these numbers the close observer might discover its spiritual counterpart, its soul. Subject this soul to close analysis and you will find

[1] The foregoing text has been largely modelled after the excellent introductory chapter in "Engineering Mathematics" by Charles P. Steinmetz.

it is the exponent which will raise some common universal base-number to the value of the number itself, and since our numbers are arranged on the decimal system, the most convenient base figure for the exponents or souls is the number ten. From this we may easily identify the souls of all *powers of ten*. Thus it is easy to see that the soul of *ten* itself is 1, since *ten* is the first power of ten. It is easy to see that the soul of *one hundred* is 2, since it is the second power of ten (that is, the base number, ten, must be taken twice as a factor to give us the number *one hundred*). It is easy to see that the soul of *one thousand* is 3; that the soul of *one million*, for example, is 6; and so on. Indeed we quickly discover that every whole or integral soul belongs to an even power of ten and coincides with the number of ciphers between the initial digit *one* and the decimal point. In short, the soul tells us the position of the decimal point.

Going back into small numbers and fractions it is equally simple. The soul of *one*, for instance, is 0, for the zero power of any number, including the base ten, is *one*. Notice that the soul still tells us the position of the decimal point, for the latter is immediately beside the initial digit, without any intervening digits. Now what is the soul of *one tenth?* Obviously it is -1, for as you know, to convert a denominator into a numerator we need merely change the sign of its exponent, and $\frac{1}{10} = 10^{-1}$. Likewise the soul of *one-hundredth* is -2, for $\frac{1}{100} = \frac{1}{10^2} = 10^{-2}$. The soul of *one-thousandth* is -3, of *one ten-thousandth* is -4, of *one-millionth* is -6, and so on. And if we write these fractions, *one-tenth*, *one-hundredth*, *one-thousandth*, *one ten-thousandth*, or *one-millionth*, as, respectively, ".1," ".01," ".001," ".000,1" or ".000,001," we shall see that their souls, namely -1, -2, -3, -4 or -6, tell us again the positions of the decimal points. Only this time, since the decimal point has been moved backwards, that is, to the left of its position beside the first significant digit (initial ciphers are not called significant), the soul has become negative. A quaint and convenient fact, that the soul always tells, by its sign and whole or integral part, the precise position of the decimal point in a number.

But what of the souls of numbers lying between the even powers of ten? There are eight numbers between *one* and *ten*, eighty-nine between *ten* and *one hundred*, and many, many more between the higher powers. They too have souls, but it is clear that they cannot have even whole or integral souls, for these belong to the powers themselves. The numbers *two*, *three*, *four* and so on, for example, lie between *one* and *ten;* so they must have souls between 0 and 1. We easily conclude, therefore, that their souls must be fractions somewhere between 0.0000 and 1.0000. That this is correct we can quickly demonstrate. Consider *the square root of ten*, a number, as you know, a little greater than *three*. It is obvious that its soul must be $\frac{1}{2}$, since the *square root of ten* is the one-half power of ten, that is $\sqrt[2]{10}=10^{\frac{1}{2}}$. Hence for the *square root of ten* we have a soul 0.5000, lying as you see between 0.0000 and 1.0000. Mathematicians have figured out to many places the souls of other numbers, that of *three* (which is but a little less than the *square root of ten*) being to four places 0.4998; that of two being 0.3010. Remember these two and you will always be able to reconstruct the souls of almost all other numbers without assistance. In short, the souls of all numbers other than powers of ten are not integral, but are fractional.

Returning to the parade, let us call a halt to the interminable thing and hold a grand review of the numbers, marshalling them according to their significant digits. In the entire battalion of numbers there will then be but nine regiments, each led by one of the significant digits, *one*, *two*, *three*, *four*, *five*, *six*, *seven*, *eight*, or *nine*. Each regiment will again be composed of ten companies in which the second digit is one of the ten numerals, *zero* to *nine*. Each company will be divided into ten platoons in which the third digit likewise varies, each platoon into ten squads whose fourth digits vary, and each squad will be similarly divided into ten subdivisions. This subdivision could proceed indefinitely. You will notice that we have here disregarded entirely the position of the decimal point. Now the interesting thing about this arrangement is the fractional parts of the souls. For the souls would never repeat themselves in this review, but there would be one fractional part of a soul assigned to each file or succession of significant digits, and belonging to that particular file al-

ways, regardless of changes in the position of the decimal point. The integral parts of the souls would indeed change with the change of the decimal point, but not the fraction. Thus glancing down the file "200000," we find that the soul of *two* ("2.00000") is 0.3010, that the soul of *twenty* ("20.0000") is 1.3010, that the soul of *two hundred* ("200.000" is 2.3010, that the soul of *two thousand* ("2,000.00") is 3.3010, and so on. In short, while the integral parts of souls are the same for similar positions of the decimal point, the fractional parts of the soul are the same for similar successions of (significant) digits in numbers.

But this glimpse of the souls of numbers must come to a close. From now on in this book (and in all other books), you will meet them again only under the prosaic names of logs or logarithms, or more precisely, common or Briggsian logarithms.[1] When logarithms are used, the natural numbers for which they stand are sometimes called anti-logarithms. The fractional part of the logarithm is called the "mantissa" and because it is the same for all similar combinations of natural numbers or significant digits, it forms the body of the table of common logarithms. The integral or whole part of the logarithm is called the "characteristic," and since it records the position of the decimal point is not shown in logarithm tables but is left to be determined by inspection. With the information dispensed in this chapter, in as heavily sugar-coated pellets as we could provide, you are prepared to meet, master and make use of any logarithm which strays your way as if it had been your life-long servant—no, no, much better than that!

In this chapter we shall go no further into the uses of logarithms. They shall sit up and perform for us through the major part of the rest of this book. We merely repeat, for your lasting remembrance (and don't ever forget it) their fundamental relations:

$$\log A + \log B = \log (A \times B)$$
$$\log A - \log B = \log (A \div B)$$

[1] "Logarithms were invented and a table published in 1614 by John Napier of Scotland; but the kind now chiefly in use proposed by his contemporary, Henry Briggs, professor of geometry in Gresham College in London."—*Century Dictionary.*

N	0	1	2	3	4	5	6	7	8	9	P. P. 1	2	3	4	5
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374	4	8	12	17	21
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755	4	8	11	15	19
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106	3	7·	10	14	17
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430	3	6	10	13	16
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732	3	6	9	12	15
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014	3·	6·	8	11	14
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279	3	5	8	11	13
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529	2	5·	7	10	12
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765	2	5·	7·	9	12
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989	2	4	7	9	11
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201	2·	4	6·	8	11
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404	2	4	6	8	10
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598	2·	4·	6·	8	10
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784	2	4·	5	7	9
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962	2·	4·	5	7·	9
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133	2	3·	5	7·	9
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298	2	3	5	7	8
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456	2·	3	5·	6	8
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609	2·	3·	5·	6	8
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757	1·	3·	4·	6·	7
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900	1	3·	4·	6·	7
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038	1·	3·	4·	6·	7
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172	1	3·	4	5	7
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302	1·	3·	4·	5·	6
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428	1	3	4·	5·	6
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551	1·	2	4	5·	6
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670	1	2·	4·	5·	6
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786	1	2·	3	5·	6
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899	1·	2·	3·	5·	6
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010	1·	2·	3·	4·	6
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117	1	2·	3·	4·	5
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222	1·	2·	3·	4·	5
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325	1·	2·	3·	4·	5
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425	1·	2·	3·	4·	5
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522	1·	2·	3·	4	5
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618	1	2·	3·	4·	5
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712	1	2·	3·	4	5
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803	1	2	3·	4·	5
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893	1	2·	3	4·	4
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981	1·	2·	3·	4·	4
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067	1·	2·	3·	3·	4
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152	1·	2·	3·	3·	4
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235	1·	2	2·	3	4
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316	1·	2·	2	3·	4
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396	1·	2·	2·	3·	4

Fig. 327. Table of Logarithms, 1-5.

N	0	1	2	3	4	5	6	7	8	9	P P 1 2 3 4 5
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474	1. 2 3 4
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551	1 2 3 4
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627	1 2 3.4
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701	1 2 3 4
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774	1 2 3 4
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846	1. .2 3.4
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917	1 2 3 4
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987	1 2 3.
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055	1 .2 3
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122	1. 2 3.
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189	1 2 3
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254	1.2 3
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319	1 2 3
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382	1 2. 3
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445	1 2 3
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506	1 2 3
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567	1.2 3
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627	1 2 3
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686	1 2 3
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745	1 2 3
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802	1 2. 3
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859	1. 2 3
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915	1 2 3
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971	1 2. 3
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025	1 2 3
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079	1 2 3
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133	1. 2 3
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186	1 2 3
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238	1. 2. 3
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289	1 2 3
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340	1 2 3
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390	1 2.3
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440	1 2
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489	1 2
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538	1 2.
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586	0 1. 2
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633	0 1 2
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680	0 1 2
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727	0 1 2
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773	0 1 2
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818	0 1 1. 2. 2
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863	0 1 1. 2. 2
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908	0 1. 1. 2 2
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952	0 1 1. 2. 2
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996	0 1 1. 2. 2

Fig. 328. Table of Logarithms, 5-9.

A device which will do all this is worth knowing.[2] In the language of valedictorians, we commend to your early attention a small table of logarithms and many pleasant hours of easier computing therewith.

---

[2] From logs, that is, the logarithms of numbers, it is but a simple step to proceed to loglogs, that is, the logarithms of logarithms. Consider the phrase A^B, in which A and B are any values we wish, such as 29.37 and 43.921. We can write $\log A^B = B \log A$. This reduces the involution to a mere matter of multiplying B into the log of A. But if the multiplication be tedious, as with the values first instanced, it will be simpler to write

$$\begin{aligned}\log\ (\log\ A^B) &= \log\ (B \log A)\\ &= \log B + \text{loglog}\ A\end{aligned}$$

and proceed by addition. The loglog of a number is the logarithm of its logarithm, and is found in the log tables by treating the logarithm as a number.

Chapter XXXII

THE LAW OF ORGANIC GROWTH

The law of organic growth, as it is called, is well-nigh as important to the practical business man as the law of cause and effect, but is unfortunately much less understood. The chart papers and methods discussed in this section of the book are designed to interpret statistics in the light of this law. To the uninitiated their construction remains a mystery, but to those who know the law which is the key to their meaning they are so valuable as to eclipse and almost to obviate all other chart methods. The law relates to the way in which a large majority of natural organic forces have been found to grow or change. It prescribes or defines the manner in which this growth or change will take place. The law is that, at regular intervals of time, each new value will be a constant percentage of the immediately preceding value.

This feature of a constant relation between successive items in a series marks what mathematicians call a progression. There are several kinds of progressions, only one of which follows the law of organic growth. By far the simplest form of progression is the one called arithmetical. In the arithmetical series or progression, each item differs from the preceding item by a constant amount (quantity, difference or increment). The series progresses from item to item either by addition or by subtraction of this amount. For example, in the series, 1, 2, 3, 4, 5, 6, . . ., the constant increment is +1. In the series 4, 3, 2, 1, 0, − 1, −2, . . ., the constant is −1. In business, the familiar instance of the arithmetical progression is the accumulation of simple interest.

Another and a very different series or progression is the one called geometrical. In a geometrical progression, each item differs from the preceding item by a constant ratio (rate, percentage, factor, multiplier, or divisor). The series progresses from item to item by multiplication or division by this con-

stant ratio. For example, in the series, 1, 2, 4, 8, 16, 32, 64, . . . ,the constant ratio or factor is 2. In the series 4, 2, 1, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, . . . the constant is $\frac{1}{2}$. In business, a familiar instance of the geometrical progression is the accumulation of compound interest. And it is the geometrical progression which the law of organic growth prescribes.

It is interesting to study these two types of progression, for they are the gist of the distinction between the curve charts which we have so far considered and the curve charts to which we are coming. In the first place let us consider the relative speed of these progressions. Compare the series 1, 2, 3, 4, 5, 6, . . ., with the series, 1, 2, 4, 8, 16, 32, 64, . . ., and you will see that the geometric progression rapidly outruns the arithmetical one. These series have begun at unity and progressed by a 100% increase, which is a fairly rapid rate of increase. But the acceleration, from the arithmetical point of view, of the geometrical series will still be evident if we take a slower rate of increase such as 10%. Starting at unity, the arithmetical series will be 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, . . . while the geometrical series will be 1, 1.1, 1.21, 1.33, 1.47, 1.62, 1.78, 1.96. . . . Of course either of the two types of progressions can begin with any item and increase at any rate, but from any point you wish to choose, if the rates are the same for the two series, the geometrical progression will always increase more rapidly than the arithmetical one.

On the other hand, in the decreasing or diminishing direction, the arithmetical progression will leave the geometrical one behind. Compare the series, 2, 1, 0, –1, –2, –3, . . . with the series 2, 1, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, . . ., and this will be evident to you. And here we come to an important distinction between the two series, namely that while the arithmetical series can reach zero and pass into negative values, the geometrical series can never reach zero at all. In the example just given, the arithmetical series diminishes by subtracting $\frac{1}{2}$ of the value of its first item and quickly passes zero, but the geometrical series diminishes by division by 2, and gives no indication of ever reaching the value of zero. Now we could have made the geometrical series begin with any other positive value and decrease at any other rate we please, but it still would be impossible for us to bring the geometrical progression down to zero. We can, by constantly diminishing it, that is, by repeatedly dividing its last item, bring it as

close to zero as we please, without ever succeeding in entirely wiping it away. In mathematical language, zero is the infinitesimal limit of the geometrical progression.

Malthus made popular the distinction between these two kinds of progression with his theory that while the population of the world increased geometrically, the wealth of the world increased only arithmetically and hence soon limited the welfare of the population. But his theory has been proved to be false, the wealth of the world appearing to increase geometrically, though sometimes at a slower rate than the population. Indeed an attempt has recently been made [1] to invest mankind with a peculiar and to some extent exclusive power of geometrical progression, both in mental and physical accomplishments, a theory to which the work of Professor Ogburn[2] in suggesting something like a geometrical progression of all that may be called human civilization, bears important confirmation. While it may not be so certain that the power of geometrical development is exclusively a property of human individuals, not shared by individual animals and plants, it is safe to say that human accomplishments, including business enterprises, are as often subject to the law of organic growth as are natural forces.

The law of organic growth therefore is the proper criterion for the business man in judging the development of a business, as it is for the economist in his study of industrial and sociological records. In applying this criterion, we must forget (for the time being at least) the amount of increase in our business from year to year and center our attention upon its rate of increase. What are the year-to-year percentages of increase? If the 1911 sales were 10% larger than the 1910 sales and the 1912 sales were again 10% greater than the 1911 sales, the business has increased in accordance with the law of organic growth. If the 1913 sales were again 10% greater than those in 1912, the increase has still followed the law. If the 1913 sales were only 8% greater than those in 1912, there has been from the point of view of the law, a definite slowing down, or falling off in the rate of growth, which must either be explained by conditions outside the control of the company, such as a general business depression, or is a harbinger of ill

[1] *Cf.* Korzybski, Alfred, *Manhood of Humanity*, E. P. Dutton & Co., New York, 1921.
[2] *Cf.* Ogburn, William Fielding, *Social Change*, B. W. Huebsch, Inc., New York, 1922.

omen which should call for almost as careful consideration by the directors as if there had been an absolute loss. On the other hand, if the 1913 sales were 15% greater than those in 1912 and the event is not to be explained by forces outside of the control of the company, there is reason for far-sighted rejoicing and thanksgiving among the directors.

That the results of the use of this criterion are radically different from the results reached by a study of the amount of change, is evident from the fact that the former may at times be directly contradictory to the latter. Let us suppose that in its first year the gross sales of the house amount to $50,000, and in the second year to $100,000, that is, there has been an increase of $50,000, or 100%. If in the third year, sales amount to $160,000, obviously the amount of increase has gone up from $50,000 to $60,000, but the rate of increase has fallen from 100% to 60%. If in the fourth year sales amount to $225,000, the amount of annual increase has again risen to $65,000 but the rate of annual increase has fallen to about 40%. If in the fifth year sales amount to $300,000, the amount of annual increase has again risen, this time to $75,000, and the rate of annual increase has again fallen, this time to 33%. This fictitious example makes clear how illusory would be any conclusion based wholly upon the amount of change from year to year and how important it is, that the annual rate of change should be watched, that is, that the records should be studied in the light of the law of organic growth.

As a matter of fact, few business houses follow closely the law of organic growth for any considerable period of time. Or perhaps it would be better to say that though operating under the law of organic growth, they fail to maintain a constant rate of change. For it is certain that they operate under this law rather than under any law of arithmetical progression. Theoretically, perhaps, given constantly similar external conditions and internal efficiency, the growth of a business house would conform to a geometrical series and illustrate perfectly the law. But as a matter of fact, the individual business house is at the mercy of a large number of external forces which lie outside of its control and do not remain constant but are ever changing, and the records of its growth therefore show a great amount of the play of what we might call chance variation. Business men are accustomed to thinking, in some fields, of a very definite saturation point in their markets. Of course when

such a point is approached, it becomes a limit which will necessitate a slowing up of the rate of increase, in spite of otherwise equal conditions, which would have favored a strict adherence to the geometrical progression. These individual variations and any approach to limiting points do not invalidate the law of organic growth, nor do they diminish its value as a criterion of business success. They are separate and additional forces imposed upon the development of the individual business, their co-action with the law of organic growth determining the fluctuating records of the house.

In entire industries, or in large aggregates of individual business records, the adherence to the law of organic growth is much closer and the operation of the law easily seen. In the last two decades, the automotive industry has afforded a spectacular illustration of a geometrical progression with a very rapid rate of increase, although a decline in this rate in recent years is terrifying manufacturers with visions of the approach of a potential saturation point in their domestic market.[3] Two other recent industries, whose entire history can be covered in the last twenty years, show similarly close adherence to the law, both the phonograph and the moving-picture industries having grown by leaps and bounds, which when analyzed in this way become surprisingly regular and uniform. The operation of the law of organic growth in business and economic affairs is even more rigid in national and world-wide records.

[3] There is another curve which sometimes fits economic data better than the logarithmic curve, namely the Gompertz curve. This is dealt with more fully later on, in the chapter on Special Projections.

CHAPTER XXXIII

RATE-OF-CHANGE ANALYSIS

To subject a business record or the events in any other phenomenon to analysis according to the law of organic growth is not really a difficult problem. It implies of course a comparison of the rate of change or ratios between each two successive items in a series, and these successive ratios are the successive quotients obtained by dividing each item by the item immediately preceding it. This method of successive divisions would, however, be tedious even for the shortest series, and when applied to the wholesale analysis of a large body of statistics, such as (in the individual business house) the records of various lines and articles, or (in economics) the statistics for many industries or sociological developments, would indeed mount up to a forbidding and costly task. Mere inspection of the data, from which we could at once detect an arithmetical progression or recognize the failure of a given series of figures to conform to an arithmetical progression, will not often suffice to dig out the geometrical progression. Indeed, a fairly close approximation to a geometrical series may be so completely veiled in the figures that it passes unnoticed through close and even expert inspection of the data. The question, therefore, is can the geometrical series be made as apparent as the arithmetical one? And how can the failure or deviation of a series of data from a straight geometrical progression be as easily measured as its deviation from an arithmetical one?

Between the two types of progressions, arithmetical and geometrical, there is a curious inter-relation which it is well to master. Let us examine again the series 1, 2, 4, 8, 16, 32, 64, . . . With the exception of the first item in this series, it is evident that all the items are merely powers of 2 and since, as you know, 1 is merely the 0 power of any number, we may include it in the series calling it the 0 power of 2.

Therefore we can rewrite this series as follows: 2^0, 2^1, 3^2, 2^3, 2^4, 2^5, 2^6. . . . Now examine these exponents and you will find that they form an arithmetical progression, 0, 1, 2, 3, 4, 5, 6. . . . In short, we find that if a geometrical series be rewritten as a series of the powers of a single quantity (the constant multiplier or rate of change), the exponents of these powers will form an arithmetical progression. This is a rule of general application and might be used as a means of defining the geometrical series.

Here we come again to logarithms. For a logarithm, you will remember, is merely the exponent by which a common or universal base figure can be raised to a given value. In rewriting the series 1, 2, 4, 8, 16, 32, 64, . . . as the series 2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, . . . we are obviously using 2 as a common or universal base for the series and the exponents, 0, 1, 2, 3, 4, 5, 6, which will raise this base figure 2, to the value of the items in the original series may be called the logarithms (to base 2) of the items in the original series. Since $1 = 2^0$, it is obvious that $0 = \log_2 1$, since $2 = 2^1$, clearly $1 = \log_2 2$, and since $64 = 2^6$, $6 = \log_2 64$. In short, in turning the geometrical series into a series of the successive powers of a common base figure, we find ourselves writing as exponents of these powers, the logarithms of the original series (to the base figure which was used as the root of these powers). And therefore we may say that the logarithms of the items in a geometrical series will form in themselves an arithmetical progression. This is indeed a very usual definition of the geometrical progression, namely that it is composed of items whose logarithms form an arithmetical progression.

Logarithms can be taken to any common base figure we desire. For example, the same geometrical progression, 1, 2, 4, 8, 16, 32, 64 . . . may be, if we wish, written as a series of the powers of 4, as follows: 4^0, $4^{1/2}$, 4^1, $4^{3/2}$, 4^2, $4^{5/2}$, 4^3. . . . Or that same series can be written as a series of the powers of any other number provided we but care to do the necessary calculating. Now as you know, logarithms are ordinarily taken to the base figure 10 and not to the base figure 2 or 4 as in the above examples. The reason for this is that our numbers are arranged upon a decimal system and by taking the base figure as 10 we are able to make the integral part of the logarithm (characteristic) a mere record of the position of the decimal point in the original number, and we are able to make

the fractional part of the logarithm (mantissa) the same for all similar succession of similar digits. Suppose therefore that we adopt 10 as the quantity whose powers we wish to substitute for the original series. In this case we rewrite the series, 1, 2, 4, 8, 16, 32, 64 . . . as the series 10^0, $10^{0.3010}$, $10^{0.6020}$, $10^{0.9030}$, $10^{1.2040}$, $10^{1.5050}$. . . . Again we find these exponents or logarithms, 0.0000, 0.3010, 0.6020, 0.9030, 1.2040, 1.5050 . . . , forming an arithmetical series.

The inter-relation between the arithmetical and geometrical series is therefore such that in a sense, both series may be spoken of as arithmetical, the former being arithmetical in its original form and the latter becoming arithmetical when its logs are used. For this reason the term logarithmic is often used synonymously with the term geometrical to distinguish the latter form of progression and its item-to-item changes, from the progression which is truly arithmetical in its original form. For though the two types of progression are made similar by the substitution of logarithms for one of them, yet it must be remembered that they are two radically different things which must never be confused with each other. In their original form or natural numbers the one progresses by addition or subtraction and the other by multiplication or division and there is a world of difference between them. That they behave similarly when logarithms are substituted for one of them, is chiefly due to the peculiar qualities of logarithms, that by their use the process of addition or subtraction may be substituted for the process of multiplication or division (instead of dividing one number into another number to get a quotient, you subtract the logarithm of the first from the logarithm of the second and the difference is the logarithm of the quotient).

In our analysis of our statistics in the light of the law of organic growth, we can find a short cut, therefore, through the use of logarithms. In other words we can turn the items in our data into their corresponding logarithms (consulting for this purpose a table of common logarithms) and then, by comparing the logarithms, we can quickly discover any uniformity or constancy in their amount of change, and so easily detect and measure the degree of adherence in the original data to the geometrical series. The deviation of the series of logarithms from an arithmetical progression with uniform amount of change, is a measure of the deviation of the original series

of data from a geometrical progression with a uniform rate of change. In practice, this method is very simple and it may be regarded as a distinct labor-saving device in the careful analysis of statistics.

In the chapter on Index Numbers, you will recall reading that relative figures (that is, percentages) can be substituted for an absolute series or series of original data. While these relatives are ordinarily computed with a single item in the series as the norm or 100%, yet they can be computed with each item taken as a percentage of the item immediately preceding it. In this case the series is called a series of chain-relatives or chain-percentages. Now it is precisely a series of chain-percentages which the method of successive divisions already mentioned gives us. But you will find that the short-cut method of successive subtraction of logarithms does not yield results in precisely the same form. For the short-cut method gives us the logarithmic differences and these differences are the logarithms of the percentages themselves. If therefore we desire to know the rate of change in terms of its percentages of change (and not in terms of the logarithms of its percentages of change) we must again consult our logarithm tables, if we are using the short-cut method, and convert the logarithmic differences back into the percentages of change (by substituting for each difference its anti-logarithm). As a matter of fact, however, the chain-relatives or chain-percentages are, except for very popular purposes, not ordinarily of sufficient importance to justify this additional labor.

You will observe, however, that we have not yet reduced the work of rate-of-change analysis of our statistics to the same simple and easy steps as are found in an amount-of-change analysis, for the use of logarithm tables and the substitution of logarithms for the natural numbers requires, even with great proficiency, a considerable amount of time and effort. And in the wholesale analysis of a large body of statistics by this method, we will not only find the work long and tedious but we should also expect to find a large number of errors creeping into the work which would be difficult to detect. The short-cut method has, it is true, eliminated the more difficult processes of division (or multiplication), and in the lack of special calculating machines, is a long step in labor saving, but the rate-of-change analysis is not yet as simple as an amount-of-

change analysis. In the next chapter this final step will be taken and through the simple use of the graphic method combined with the use of logarithms, the work of substituting logs for natural numbers will be eliminated and the rate-of-change analysis made as simple as the amount-of-change analysis.

Organic, percentage, geometric or logarithmic change (whichever name you prefer) is growth in which the rate or ratio of change is uniform. Increment, difference or arithmetic change is growth in which the amount or quantity of change is uniform. The former naturally forms the basis of judgment for the fluctuations of phenomena which cannot be negative, that is, which must always be positive. The latter is frequently the better basis of judgment for the fluctuations of phenomena which can be zero and negative as well as positive. In general it is perhaps best to study your data from both points of view.[1]

---

[1] Speaking of the amount-of-change curve, Professor Marshall says: "Its defects are such that many statisticians seldom use it except for the purpose of popular exposition, and for this purpose, I must confess, it has great dangers."—Alfred Marshall, *On the Graphic Method of Statistics*, Jubilee Volume of the Royal Statistical Society, June 22–24, 1885, pp. 251–260.

CHAPTER XXXIV

RATE-OF-CHANGE SCALES

The rate-of-change curve chart affords in some respects the most powerful analysis known of statistical data. An attempt has been made in the last chapter to explain the general theory of this chart method, but a real insight into its various uses can only be obtained from a study of its applications. The method is really nothing more than the charting of the logarithms of numbers in the place of charting the numbers themselves. A careful reading of the last chapter will doubtless have already suggested this process to the student, and it only remains to set forth the technique of charting logarithms. Indeed, as will be seen, such simplified methods have been developed that it is not necessary for one to understand logarithms or be proficient in their use in order to benefit from this chart. In the present chapter the development and construction of these simplified charts will accordingly be discussed with the general principles covering the use of logarithms in the charts.

Three methods are open to us in the plotting of logarithmic curves. The first is the obvious one of substituting for the items in a series to be plotted, the logarithms of those items. We must consult a table of logarithms and for each item find the logarithm and tabulate these logarithms in a column beside our original series of data. Then on the plain co-ordinate paper used for amount-of-change curves, in which the scales are arithmetically projected (that is, the scale-figures at equal distances form an arithmetical series), we must plot these logarithms, and draw the curve through these plotted points. The result, of course, will be a curve of the logarithms of our original series, or, as we have called it, a logarithmic (or "rate-of-change") curve. This curve will behave as a logarithmic curve should and will tell us what we wish to learn from the use of logarithms. The straight line, which always

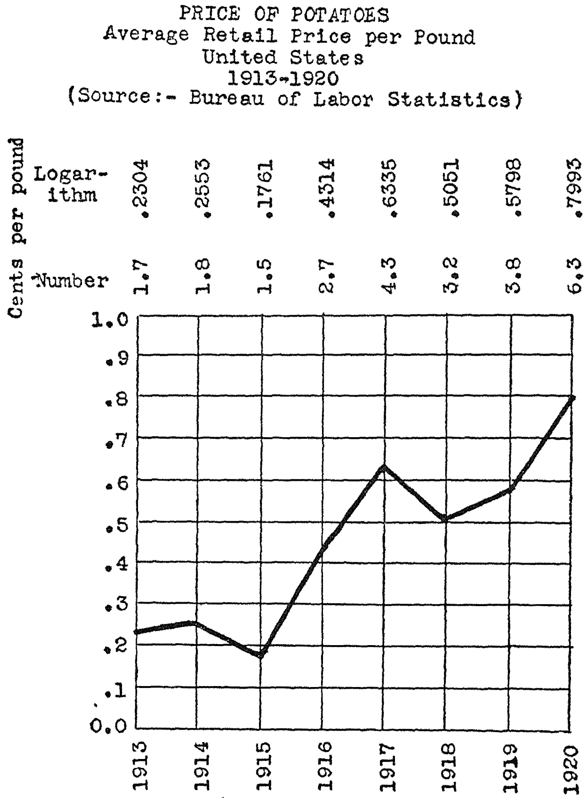


Fig. 329. The Rate-of-change Curve—First Method.
This scale carries the logarithms, not the numbers

represents equal amounts of change and therefore depicts an arithmetical progression, will here indicate an arithmetical series of logarithmic values and hence a geometric series in the original data—thereby instantly betraying to us the fact that our phenomenon has for the length of the straight line followed the law of organic growth. And the failure of our curve to maintain a straight line will indicate the failure of our phenomenon to follow the law of organic growth. All this is as it should be, but the method of charting is tedious.

The other two methods open to us achieve precisely the same resulting curve on the chart, but obviate the need of turning our original figures into logarithms. No need to bother with a table of logarithms, nor indeed, to understand the so-called intricacies of such a table. The trick is turned by

merely converting the scale of the chart, once and for all, beforehand, into a logarithmic scale. That is to say, we must calibrate the scale figures for the natural numbers, but enter these calibrations or scale figures at points on the scale which are plotted, graduated, or measured, at the values not of the natural numbers themselves, but of their logarithms. Such a scale we shall throughout the remainder of this book call a logarithmically projected scale.

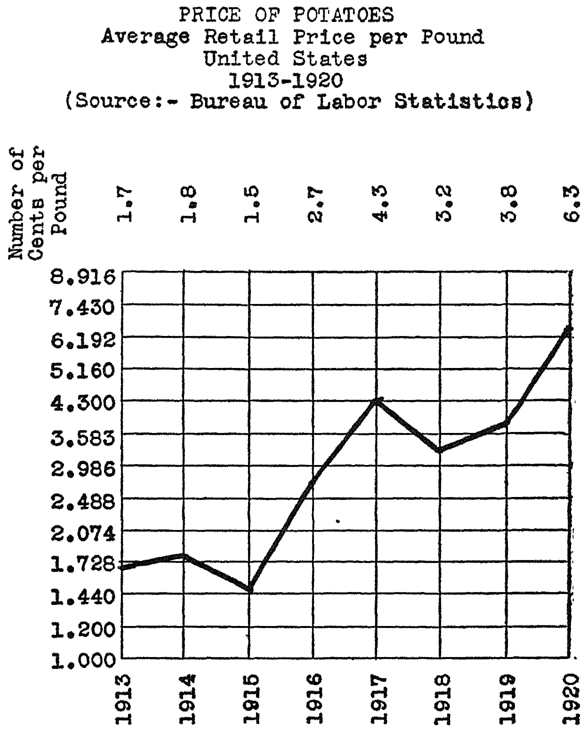


Fig. 330. The Rate-of-change Curve—Second Method.
This scale carries the numbers, not the logarithms, but is not handy.

The first of these two simpler methods uses the same arithmetically ruled co-ordinates which we have used for amount-of-change curves.[1] The scale is therefore somewhat unhandy. For if every equal interval or distance up the paper or scale is to stand for an equal logarithmic value, it must stand for an equal arithmetic or natural number ratio. If we

[1] For a full description of this method, see Irving Fisher, *The Ratio Chart for Plotting Statistics*, American Statistical Association Quarterly, June, 1917, p. 578.

calibrate the first (i.e. lowest) abscissa (horizontal line) as unity or 1.0, and let each distance or interval between the horizontal lines stand for a 10% increase (that is ratio of $\frac{11}{10}$), then obviously we must calibrate the second horizontal as 1.1, the third as 1.21 (that is $\frac{11}{10}$ of 1.1), the fourth as 1.331 (that is $\frac{11}{10}$ of 1.21), the fifth is 1.474, the sixth as 1.622, and so on. This is what we would call an unhandy scale. It is difficult to plot points on such a scale. Nevertheless, it can be done, and the resulting curve will be the same as secured by the previous method of plotting the logarithms of our series. And we have avoided the task of turning each figure of the series individually into a logarithm. And by either method you will notice that we have been free to make our curve fluctuations as high or as low as we wished, by merely selecting our scale on a larger or smaller unit length.

The third method is, however, the best of all, for it makes the plotting of logarithmic curves as simple and easy as the plotting of arithmetical ones. It consists in using specially ruled paper, provided by many publishers of chart paper, in which the co-ordinates are unevenly spaced so as to correspond with the logarithmic values of the round numbers in the original series. Thus instead of an abscissa or ordinate at the value of 1.21 (equidistant with the abscissa or ordinate of 1.0 from the abscissa or ordinate of 1.1), this paper has the abscissa or ordinate of 1.2, slightly closer to that of 1.1. Likewise instead of an abscissa or ordinate for 1.331 (at another equal distance), this paper has the abscissa or ordinate of 1.3 still closer to that of 1.2. So it goes throughout the scale. The paper has been carefully ruled up with these gradually diminishing distances or intervals between ordinates accurately measured to correspond with the true logarithmic distances of the round numbers from 1 to 10 and all fractions between these round numbers.

And since, as you have seen, the logarithms of every similar succession of significant digits are the same (in mantissa), we need merely multiply or divide these round numbers in the printed scale of this chart-paper, by any power of ten to make the scale suitable for our data. This is the same as saying that we can shift the decimal point as far in either direction of these printed scale figures as we please, and the paper will still be properly ruled off and scaled. Again it is the same as saying that we may add or prefix as many ciphers as we want to these

printed scale figures. The changing of the printed scale-figures running from 1 to 10 into a scale in which the round figures

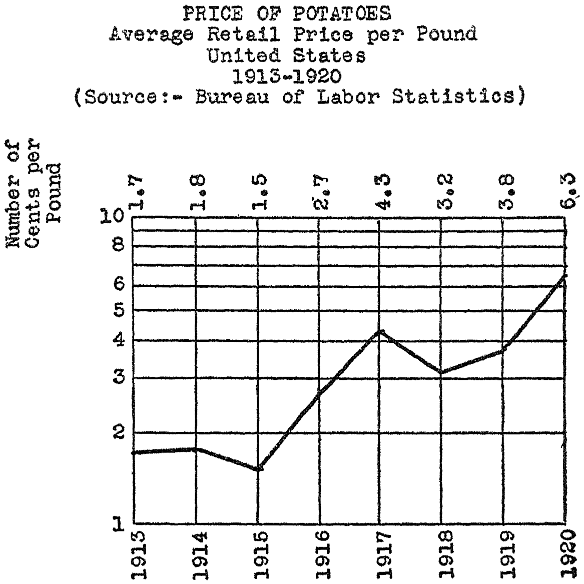


Fig. 331. The Rate-of-change Curve—Third Method.
The simplified and handy scale of original numbers.

fit our data is very easy. The only thing to remember is that it is done by multiplying or dividing the printed scale figures by a constant—whatever constant we please. In this it differs from the changing of scales on the amount-of-change curves—in which we could have used addition and subtraction. The writing in of ciphers behind or before the printed scale-figures, is merely a form of multiplication or division, in which the constant is some power of ten. It is indeed perfectly possible to use for our constant some figure which is not a power of ten. But we must multiply or divide by this constant—we cannot add or subtract.

Now it often happens that a scale running from 1 to 10, that is in which the maximum of the scale is ten times the minimum, does not afford us sufficient range for the fluctuations of our data. Were we to attempt to plot our curve upon this specially prepared paper, we would find that the curve would quickly run off the chart. To this problem the answer is very simple. We merely join together two sheets of this

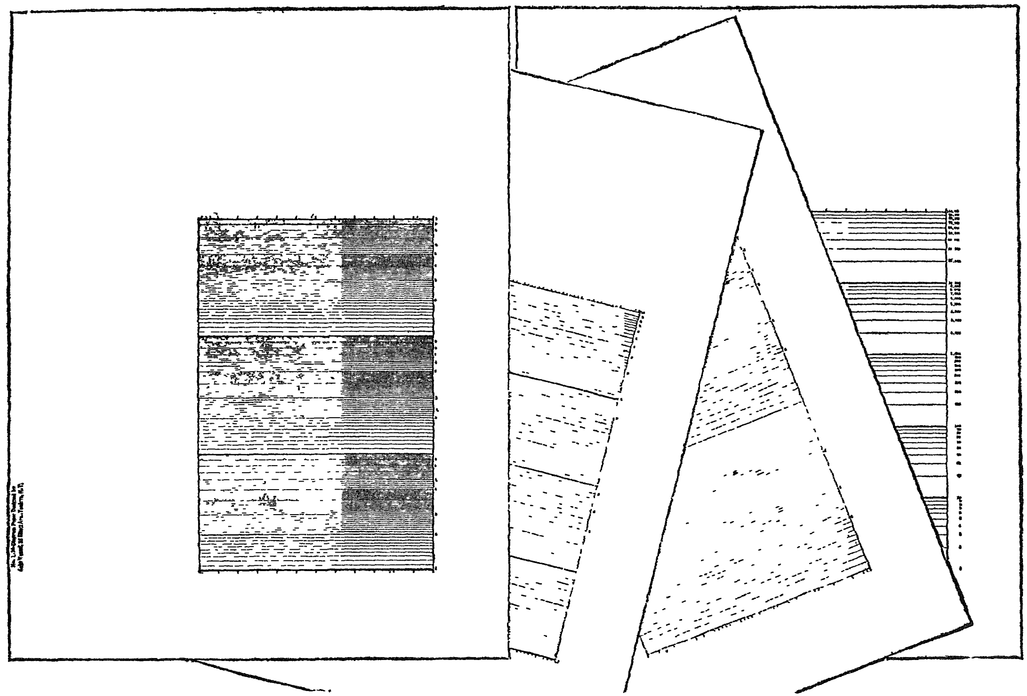


Fig. 332. Good Rate-of-change Chart-Fields.
Useful for statistical work, as space is allowed for data above the charts. The forms are on 8½ x 11 inch pages, here outlined with light lines. The forms comprise 2-deck, 3-deck, 4-deck and 5-deck paper.
—*Published by Mr. John Wenzel, Yonkers, N. Y.*

paper, or two sets of these rulings, and recalibrate the upper one by adding an extra cipher to its scale figures. In this case, the entire scale, through two sets of rulings, runs from 1 to 100—that is, the first runs from 1 to 10 and the second runs from 10 to 100. Simple, isn't it? Not only two, but many of these sets of rulings, can be joined together in this way, giving us a scale range of from 1 to 1000, 10,000, 100,000, 1,000,000, or more. In fact, the publishers of chart paper have anticipated this need, and provide paper with these sets of rulings combined two, three, four and sometimes more, upon a single sheet. Each set of rulings is called a "deck." The single-deck paper runs from 1 to 10, the double-deck from 1 to 100, and so on, two and three-deck papers being the most generally useful ones.

These considerations of the number of decks needed for a chart are all based upon the range of fluctuation in the series to be plotted. Before determining upon the number of decks to use in your chart, you must first glance through your series and note not merely its highest but also its lowest items. If both have the same number of (integral) digits, a single deck is sufficient; but if the maximum has more (integral) digits than the minimum, you need as many additional decks as there are additional digits in the maximum figure. A little thought will enable you to determine exactly the number of decks you need before you start your chart.

A very different problem is the size of decks used in your chart. In order to make their chart-forms of uniform over-all size, the publishers of this paper are accustomed to make the decks smaller as they join more of them together. Thus if in single-deck paper the chart measures six inches to the deck, in double-deck paper the deck will measure three inches that the chart may still measure six, and triple-deck paper will have three two-inch decks and in four-deck paper the four decks will each measure one and one-half inches. This, for general appearance, is excellent, but you must not mix the various sizes of decks in the same report or set of charts. You must maintain a uniform size of deck, for a chart upon a three-inch deck cannot be compared with one upon a two-inch deck. The deck on the smaller scale will show the same fluctuations smaller than they would be upon a deck on a larger scale. So if you use, let us say, the two deck paper in which each deck measures three inches, you must keep on using that paper so long

as you are charting statistics to be compared with it, regardless of whether your data at times calls for only one deck or for three or more decks (in this latter case you must make up three or more deck paper to fit the two deck paper, the decks being uniform and the chart larger). So it is best, therefore,

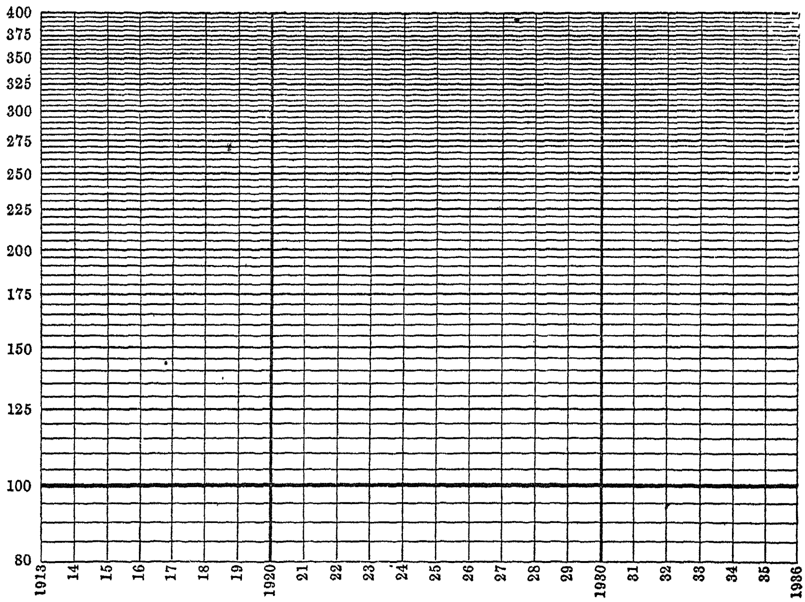


Fig. 333. A Rate-of-Change Part-Deck Form.

This form, designed by the author, for price-fluctuations, is used by the Bureau of Labor Statistics both as an office-form and in its publications on prices. The horizontal faint-rulings (those without scale-numbers) are blue in the original, so as to disappear in reduced reproductions.

before preparing a series of charts, to inspect all your data and pick once for all the best paper suitable for the most widely fluctuating series in the data.

In addition to the one, two, and three or more deck papers, you may occasionally need, and can also obtain from some publishers part-deck paper. This is useful for showing fluctuations which do not cover a range of more than 1 to 3 or 4, that is, in which the maximum item is not more than 200 or 300 per cent greater than the minimum item. Here you may find that a whole deck would waste paper and fail to show the fluctuations clearly enough. You may therefore feel called upon to adopt a chart-ruling covering only a part of a deck. But, as

will be later shown, it does not pay ordinarily to carry this detail too far, for as you take a smaller and smaller part of the deck you will find the rulings approaching nearer and nearer to plain arithmetical or uniform distances and your curve resembles more and more closely a plain amount-of-change curve.

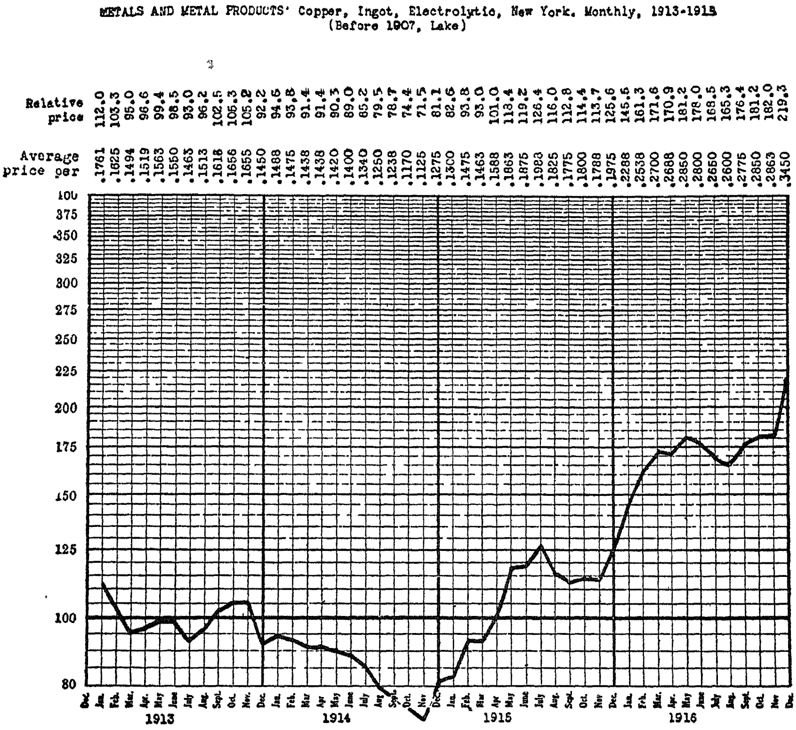


Fig. 334. Part-Deck Rate-of-Change Paper.

The curve-chart record form used in the Bureau of Labor Statistics for price relatives. This is an office form—in this case filled out for electrolytic ingot copper prices at New York—the chart nearly filling 8½ x 11 inch paper sheets.

There is, however, still another form of ruling which, you will often find useful, known as the "split-deck." By means of a split-deck you can often keep to larger decks though your maximum and minimum figures in a series lie in additional decks. The split-deck is merely the upper half of one deck and the lower half of another joined together in a single chart. You may have split one-deck paper or split two-deck

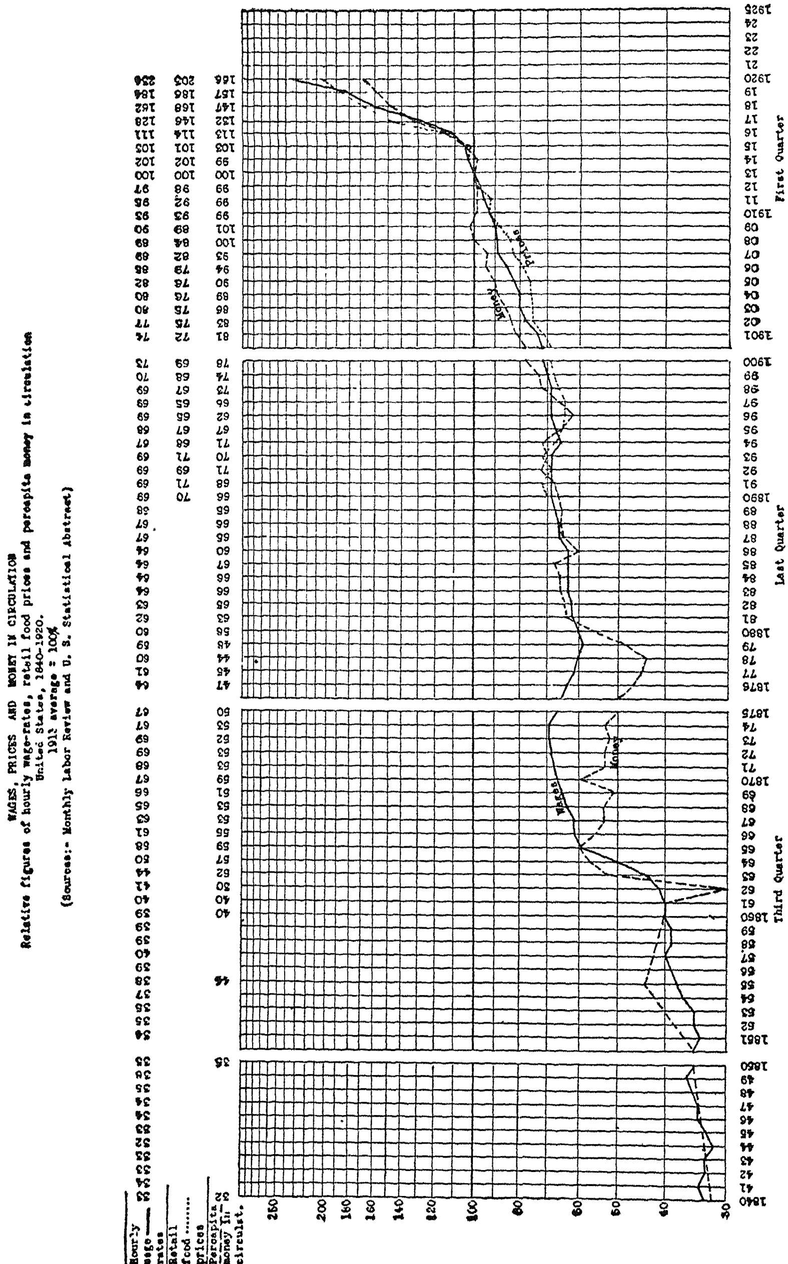


Fig. 335. A "Split-Deck" from 30 to 300.

It is easier to convert a single deck to 20-200 or 50-500. This 30-300 scale was specially drawn, from a slide-rule.

paper, or more, but the first is the only kind ordinarily published. By means of a split-deck chart-field you can often keep to a larger size of deck and yet show a curve whose maxima and minima lie in different decks. Any paper can be converted into split-deck paper merely by multiplying the printed scale figures by some constant other than ten or a power of ten—two and five being the usual and most convenient constants for this purpose.

If you do not have access to the marketed forms of specially ruled logarithmic chart-paper, you can readily prepare it for yourself, either by plotting for your scales along the axes, the logarithms of the round numbers, as found by consulting a table of logarithms, or by copying the calibrations and graduations from an ordinary slide-rule. If the slide-rule (in which the deck is usually about five or ten inches long) does not have the right size of deck to suit you, and in general if you wish to alter a scale as to size of decks, you can accomplish this by the trick of laying off (or "projecting") the given calibrations to precisely the size you desire by the use of parallel lines from the given scale to the desired scale, across a triangle formed by the two scales and the last parallel.[2]

It is a pretty way of ornamenting the page on which a rate-of-change chart is shown, to mark off a short additional scale near, but not as part of, the chart. This scale need not be as long as the scale upon the chart. It may be made up of two parallel lines close together, with cross-lines at the round numbers on the scale, the whole looking very like a long narrow ladder. If you wish to make it more conspicuous, the alternate spaces between cross-lines can be blacked in. The virtue of this gratuitous and emphasized scale is that it calls the layman's attention to the strange and unusual (to him) nature of the scale of the chart. And its strong markings also show the constant significance of distances upon the chart, regardless of height. By recalibrating this extra scale with scale-numbers of "per cent increase or decrease" (that is, 0 at the point of 100, 50 at the point of 150,-25 at the point of 75) you make this constant significance of distances even clearer. The small extra scale is then an excellent device for use with a pair of dividers—it is like a scale of miles on a map. The reader

---

[2] Cf. Figs. 166 and 167 on p. 185.

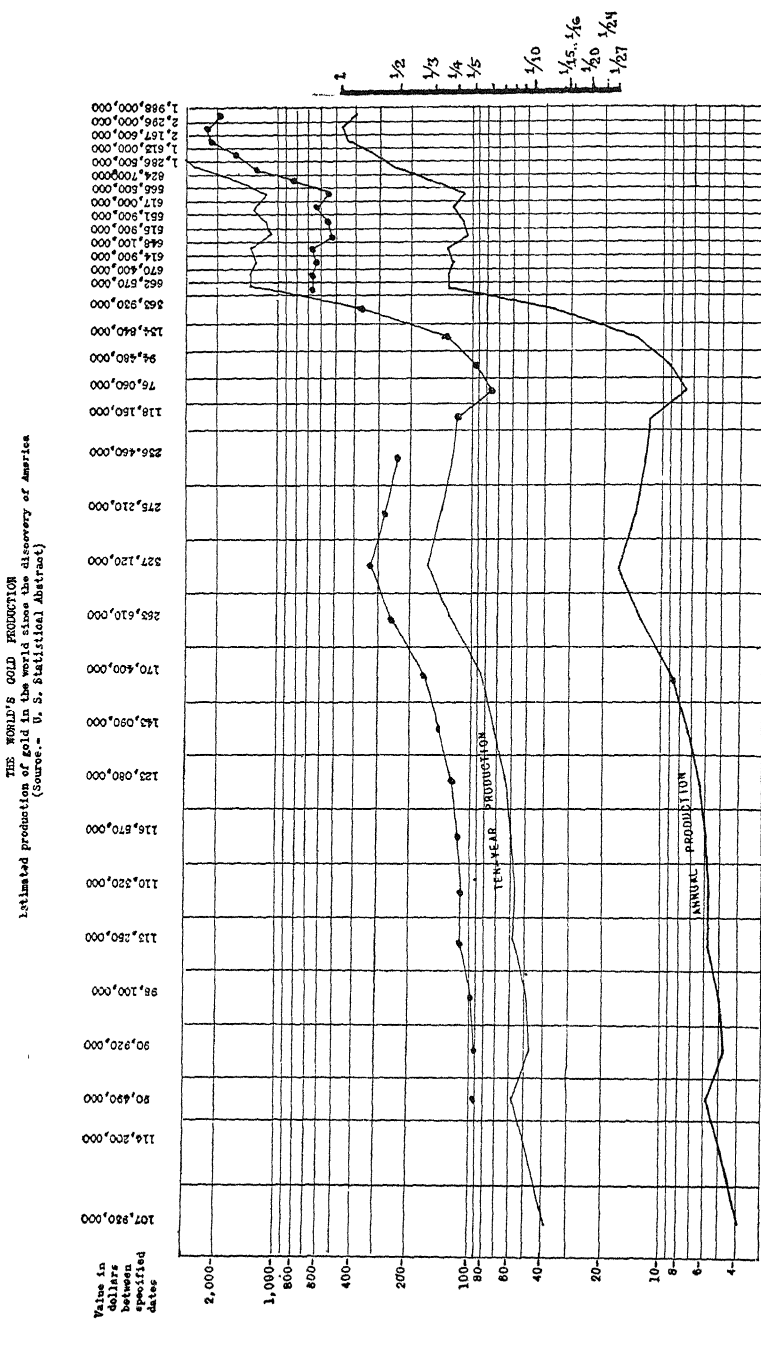
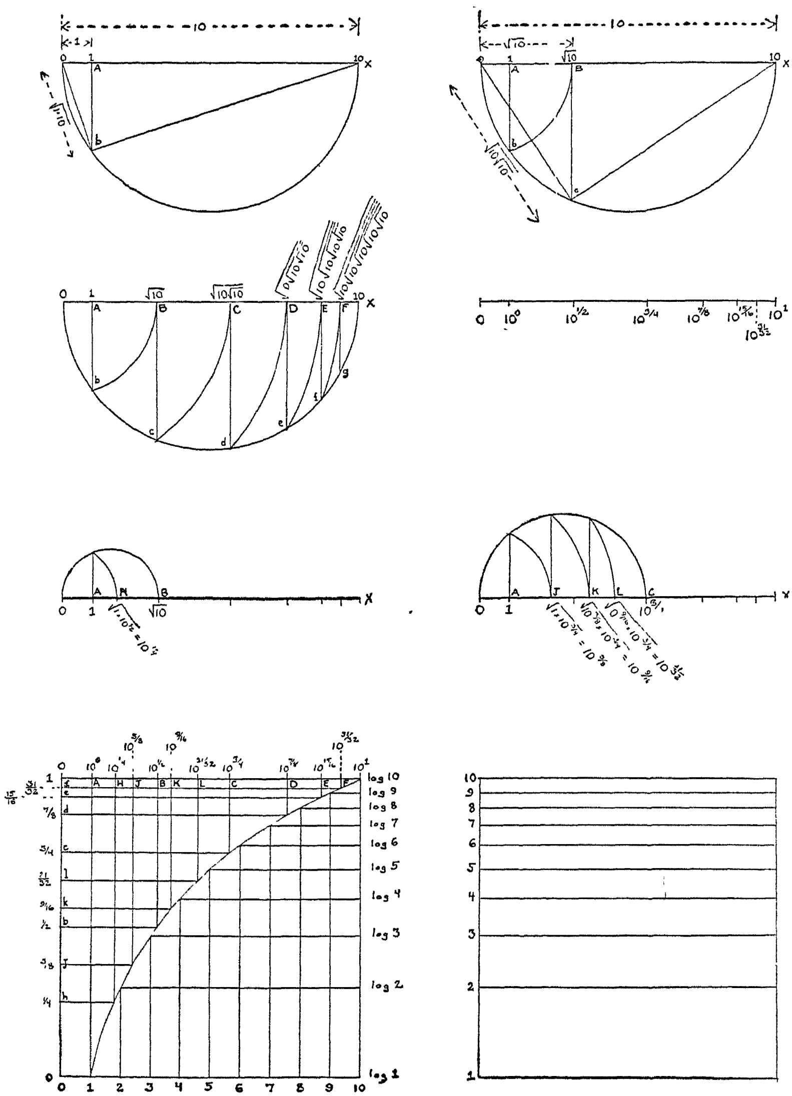


Fig. 336. Ncte the Special Scale at the Right (See Footnote on opposite Page).

can adjust his dividers for any two points, hold the dividers to the scale, and read the percentage relation between them. And this, after all, is the purpose of rate-of-change charts.

NOTE TO FIG. 336

Fig. 336 illustrates the computing possibilities of rate-of-change paper. The data, as will be seen above the chart, is for periods of irregular length as shown by the dates below. While the points have been plotted for the data, these points cannot all be connected as a single curve. It is necessary, as pointed out in the chapter on frequency curves, to find averages for periods of uniform length. We need not, however, calculate these averages—on the contrary with a pair of calipers or dividers we lay off the proper distances below the points from the special scale at the right (or, in its absence, from the regular scale at the left) and obtain at once the plotting points for a single connected curve. In this chart two such calculated curves have been drawn—one for an annual average and the other, just one deck higher, for a decennial average.



10
X
A
b
B
C
D
E
F
log 10
log 9
log 8
log 7
log 6
log 5
log 4
log 3
log 2
log 1
0 1 2 3 4 5 6 7 8 9 10

Fig. 337. One Way to Find the Rate-of-change Scale.

On any horizontal distance OX lay off OA equal to one-tenth of OX and project a semi-circle on OX (see upper left-hand diagram); drop a perpendicular Ab from OX to intersect the semi-circle at b. Then $OA : Ob :: Ob : OX$ and $Ob^2 = (OA)(OX) = (1)(10)$ and $Ob = \sqrt{10}$. Lay off OB on OX equal to Ob (see upper right hand diagram) and drop a perpendicular Bc from OX to intersect the semi-circle at c. Then $OB : Oc :: Oc : OX$ and $Oc^2 = (OB)(OX) = 10\sqrt{10}$ and $Oc = \sqrt{10\sqrt{10}}$. Likewise (see third diagram) lay off OC, OD, OE, and OF on OX. These distances (see fourth diagram) represent $\sqrt{10\sqrt{10}}$, $\sqrt{10\sqrt{10\sqrt{10}}}$, $\sqrt{10\sqrt{10\sqrt{10\sqrt{10}}}}$ and $\sqrt{10\sqrt{10\sqrt{10\sqrt{10\sqrt{10}}}}}$ respectively, or $10^{3/4}$, $10^{7/8}$, $10^{15/16}$, and $10^{31/32}$, while we already have in OA, OB, and OX, obviously 10^0, $10^{1/2}$, and 10^1. Similarly (see fifth and sixth diagrams) we can find $10^{1/4}$, and $10^{3/8}$, $10^{9/16}$, $10^{21/32}$ from semicircles upon OB and OC. The ordinates from these points (see lower left-hand diagram) intersect abscissae from a scale of the exponents, 0, $\frac{1}{4}$, $\frac{3}{8}$, $\frac{1}{2}$, etc. so as to form a logarithmic curve, or curve of the logarithms of numbers, in which $y = \log x$. By taking abscissae from the intersect points of this curve with ordinates from the even numbers from 0—10, we find the logarithmic scale of these numbers (see lower right-hand diagram).

Chapter XXXV

RATE-OF-CHANGE CURVES

From the construction of the logarithmic chart, we turn naturally to the general principles of its application, and here we must consider two of its limitations. The first of these has already been touched upon, that the utility of the logarithmic projection is limited to data in which the range of fluctuation is fairly great. There is not much to be gained from the logarithmic projection when the maximum of a series does not vary by more than a 100% from the minimum of the series. For the logarithmic scale and the arithmetic scale approach each other more and more closely as the percentage between the limits is made smaller and smaller. Within a range of 100%, that is, when the maximum or upper limit is not more than twice as great as the minimum or lower range, the difference between the two projections or scales is not great enough to justify ordinarily the effort of the less usual method. Within a range of 50% variation the approximation of the two projections or scales is very close indeed, and within a range of 25% variation, the difference is hardly noticeable. The real value of the logarithmic projection is for data which fluctuates to ranges exceeding two or three hundred per cent or more.

This fact, that when only slight changes take place in the total series under observation the geometric progression closely approximates the arithmetical one, enables us to use the latter, because of its simplicity, even for the purposes of interpolation and extrapolation. Thus while the population of the United States increases by about two per cent per annum, a geometrical progression, yet the Census Bureau employs constant differences or amounts of change in estimating the local population for the inter-censal years. The results are sufficiently reliable, because at two per cent it would take the geometrical progression many years to outdistance the arithmetical progression noticeably. And in the charts which fol-

low, designed to show geometric progressions instead of arithmetical ones, it must be remembered that no great benefit is secured from charts in which the range of fluctuation of curves is slight.

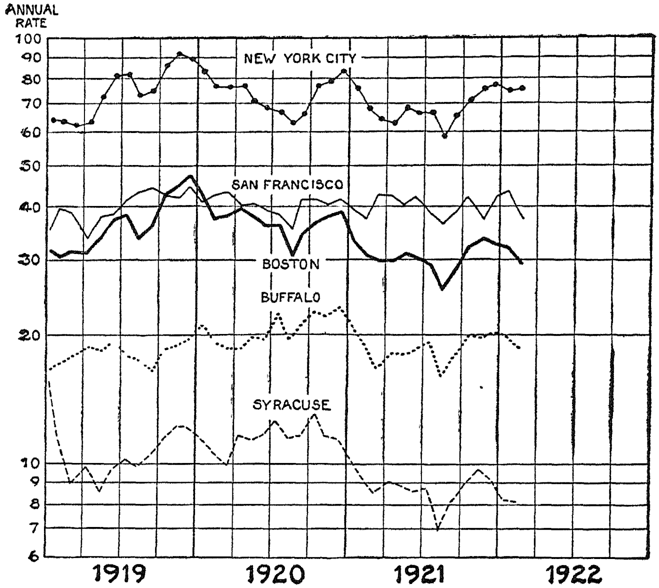


Fig. 338. Comparison of Series Lying in Different Parts of Chart, Though Not Fluctuating Greatly.

Annual rate of turnover of bank deposits in representative groups of banks in different cities.—*Permission of Mr. Carl Snyder.*

To this limitation of the usefulness of the logarithmic chart method, we must note an important exception, arising when a number of different series are to be compared and although their individual fluctuations are slight, yet they would lie upon far different portions of the chart. In the chapter on Index Numbers, you saw that such quite different series could be easily compared by reducing them to percentages, the change into percentages or relatives making their fluctuations comparable when charted upon arithmetical or ordinary amount-of-change chart paper. The use of the logarithmically-projected chart scale, that is, the rate-of-change paper, will however obviate the need for mathematical computing required to change the series into relative percentages, as the logarithmic projection makes their fluctuations comparable regardless of their position upon the field of the chart. Thus

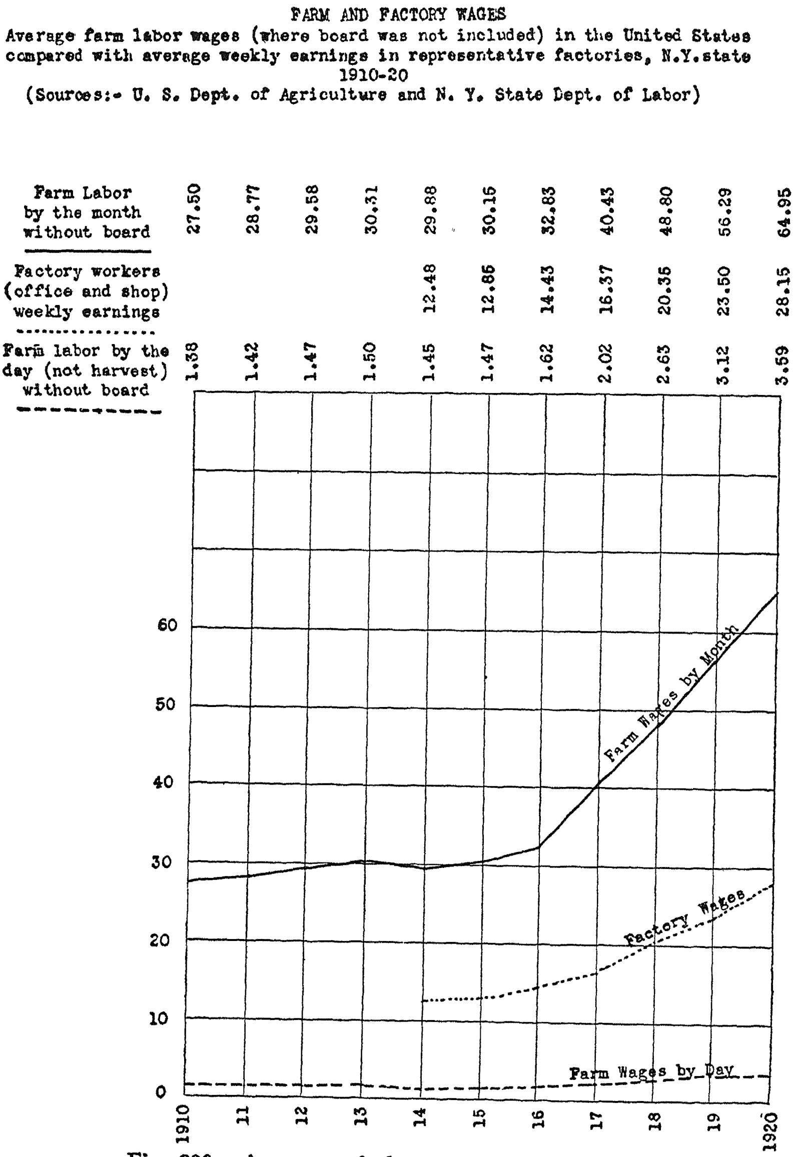


Fig. 339. Amount-of-change Curve (**Absolute**).

two series of data which fluctuate similarly as to rate of change, appear very unlike when plotted arithmetically, if one lies further from the zero or base line than the other (as when the units of measurement of one are millions of dollars, and of the

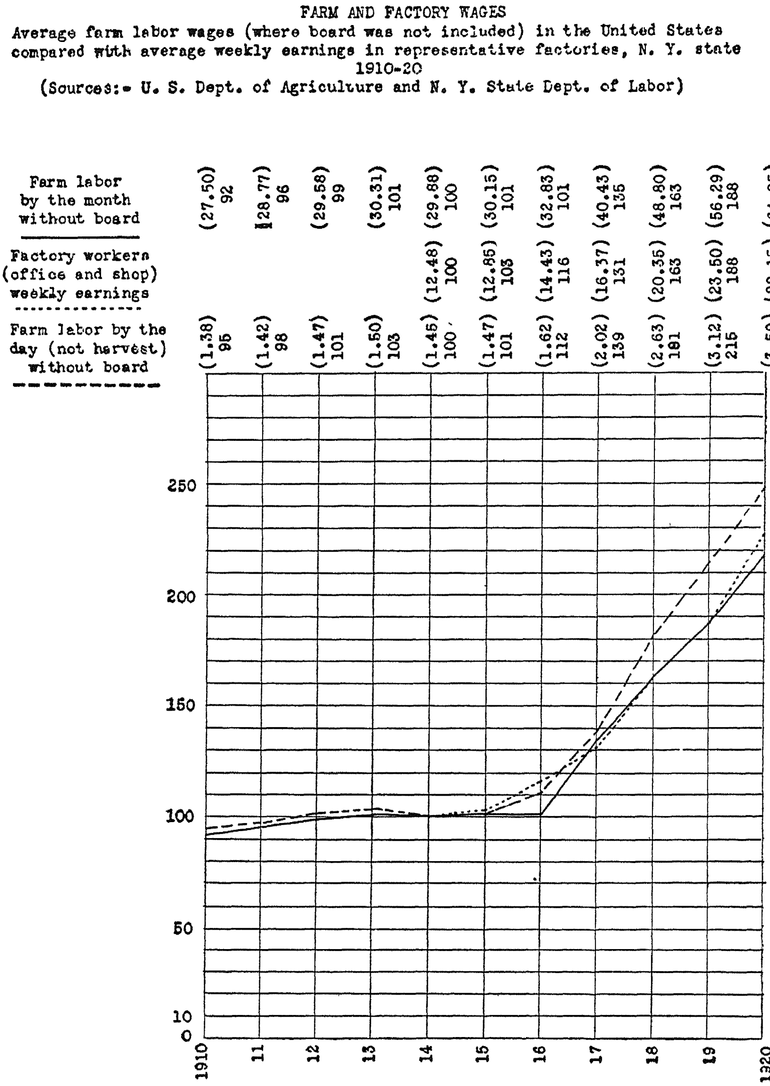


Fig. 340. Amount-of-change Curve (Relative Numbers).

other dollars and cents). But when logarithmic paper is used, these two series appear to fluctuate together (just as they did when reduced to index numbers or relatives). The logarithmic chart method moreover avoids all confusion which might arise as to the base year or period employed for the two relative series. This advantage becomes important when different

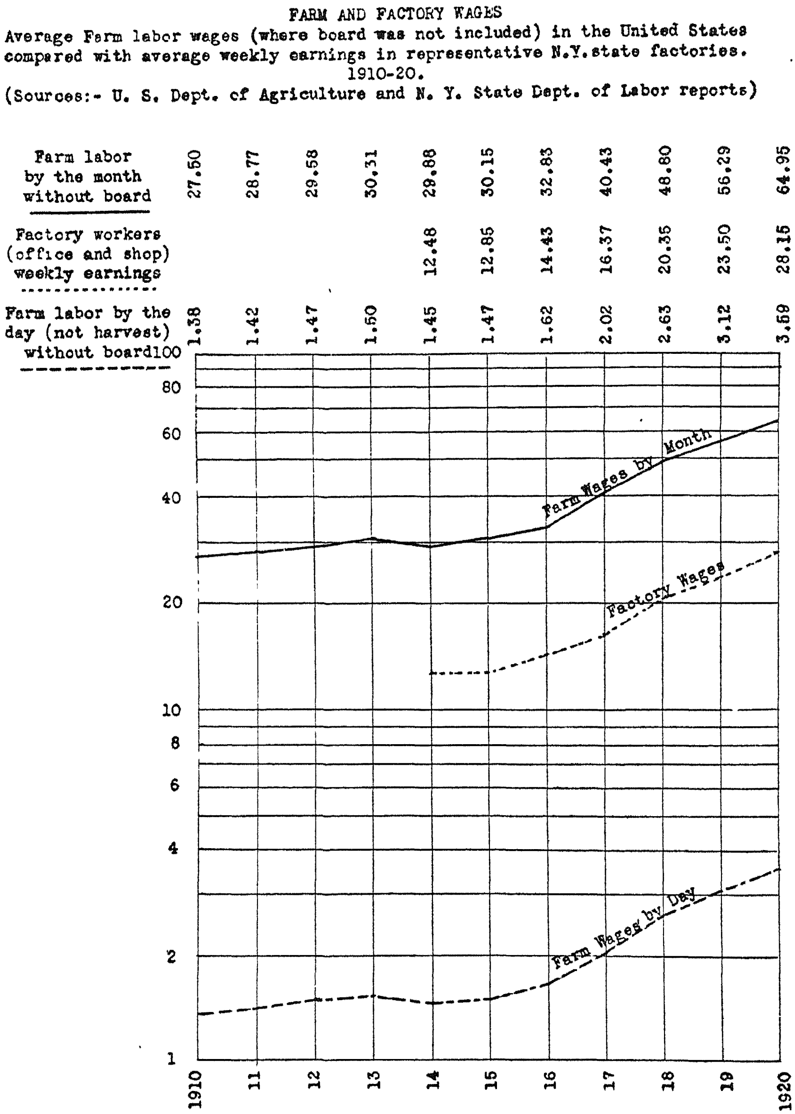


Fig. 341. Rate-of-change Curve (Absolute).

periods of time are used as the bases for the two relative series, for even identical series would differ from each other with different base figures when plotted upon amount-of-change paper.

The other limitation of the logarithmic or rate-of-change chart (and a much more important limitation) is that this

chart can be used only for values in which zero is an absolute limit. The chart cannot be used for data in which the values cross from positive into negative numbers or vice versa. Zero is an absolute limit to any geometric progression and to anything operating under the law of organic growth. The population of a community can never be zero (if it is to remain a population) and it cannot be a negative quantity. The production of a factory might be zero, but it cannot be a negative quantity. The sales of a concern cannot be negative. Innumerable examples could be given of data to which zero is an absolute limit and to all such data the rate-of-change chart is not only applicable but proper.

To the rule that zero values cannot be shown upon a logarithmic chart, there is an apparent exception, to be found in the case of data measured in units which are made upon an arbitrary and not upon an actual, zero-point. The Fahrenheit scale of temperature, for example, places its "zero" value a short distance below the freezing point for water at sea level. Zero here is an entirely arbitrary valuation and does not mean an actual nothing, that is, it does not mean zero heat. To plot temperature by taking the logarithm of Fahrenheit degrees themselves would be ridiculous. Not only would we be unable to show zero degrees Fahrenheit (because the logarithmic scale cannot reach zero) but indeed the shape of any curve which we might plot in this way would be meaningless. The sensible procedure would be to plot the temperature after changing the readings into the absolute scale of temperature, or more simply, to prepare a special scale in which the degrees were plotted at their values on the absolute scale. This is done by graduating the scale according to the logarithms of the absolute degrees of temperature, and then recalibrating or labelling these graduations with the equivalent Fahrenheit readings. After this recalibration or special labelling the figure O would of course appear upon the scale of the logarithmic chart, having been entered at the scale-point which really represented about 265, the point on the absolute scale corresponding to 0° F. This example makes clear that a zero reading or value can appear upon a logarithmic chart when it is fictitious and really represents a positive value. Another example of the same type would be a scale of time which included the year 0 A. D., from which we date our years in modern history. All such are cases in which the real values plotted upon the chart

have actual positive values, which through some peculiar circumstances must be assigned zero or negative values to conform to ordinary practice.

A more general example of this recalibration of the scale resulting in zero and negative values is to be found in percentage scales in which the hundred per cent point or line has been relabelled zero and all other figures on the scale correspondingly relabelled to represent percentage of increase or decrease from this particular point. What has really happened here of course is that 100% has been subtracted from every point along the scale in order to get the desired calibration.

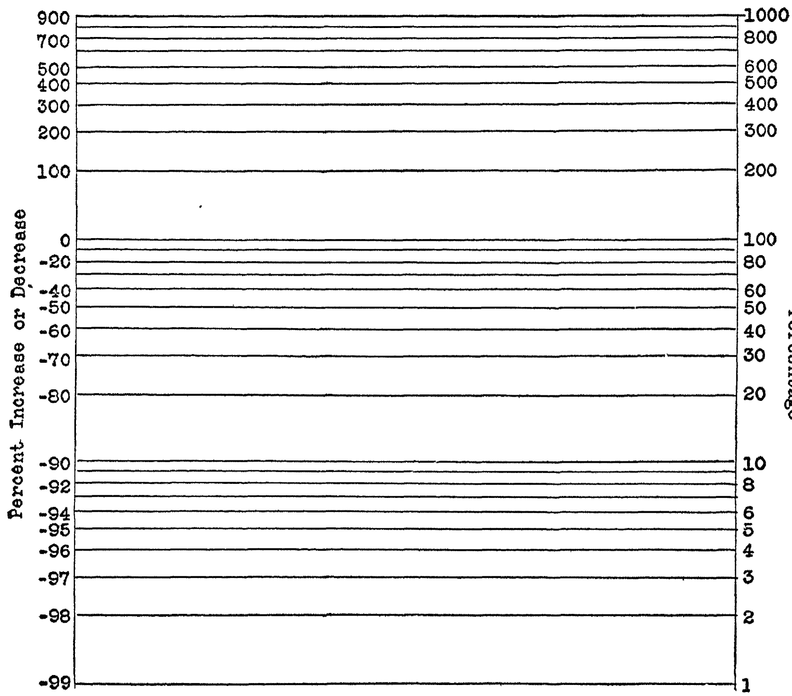


Fig. 342. The Percentage-Increase-or-Decrease Recalibration.

It is an apparent (though not a real) exception to the rule given in an earlier paragraph on the construction of logarithmic charts in which scale changes were said to be made only by multiplication or division of the scale figures, and not by addition or subtraction. Reading upward on this new "percentage increase or decrease" scale from "0" (entered at the true point of 100), we find "+50" entered at the true point of

150, "+100" at the true point of 200, "+900" entered at the true point of +1000%, and so on. Reading downward we find "−10%" entered at the true point of 90%, "−50%" entered at the true point of 50%, "−90%" entered at the true point of 10%, and so on, the chart approaching but never reaching the point of −100%, which belongs to the real zero value which cannot be shown upon the logarithmic projection. Such a recalibration as this special "percentage increase or

ACCIDENT MORTALITY RATES
(Mortality Rates per 1,000 Population of Specified Accidents)
United States
1910-1912
(Source:- United States Bureau of Census)

Males	Falls	.096	.054	.041	.054	.097	.129	.160	.191	.234	.280	.327	.366	.457	.552	.879	1.3[illegible]3	2.834	5.39[illegible]	8.608	9.52[illegible]	1.495
	Drowning	.077	.132	.151	.197	.207	.168	.160	.171	.181	.177	.176	.154	.132	.143	.117	.080	.150	.160	.072	....	.374
	Railways	.014	.037	.062	.163	.338	.372	.379	.414	.387	.398	.361	.369	.377	.342	.344	.434	.389	.377	.398	.374	....
Females	Falls	.064	.022	.012	.010	.012	.015	.020	.031	.037	.046	.081	.136	.220	.462	1.013	2.357	5.091	9.811	16.230	14.463	4.504
	Drowning	.037	.016	.019	.026	.022	.012	.011	.011	.014	.014	.010	.011	.010	.011	.016	.027	.026	.016	....	....	....
	Railways	.009	.009	.009	.012	.014	.014	.015	.017	.021	.022	.029	.033	.041	.044	.073	.090	.069	.079	.085	....	....

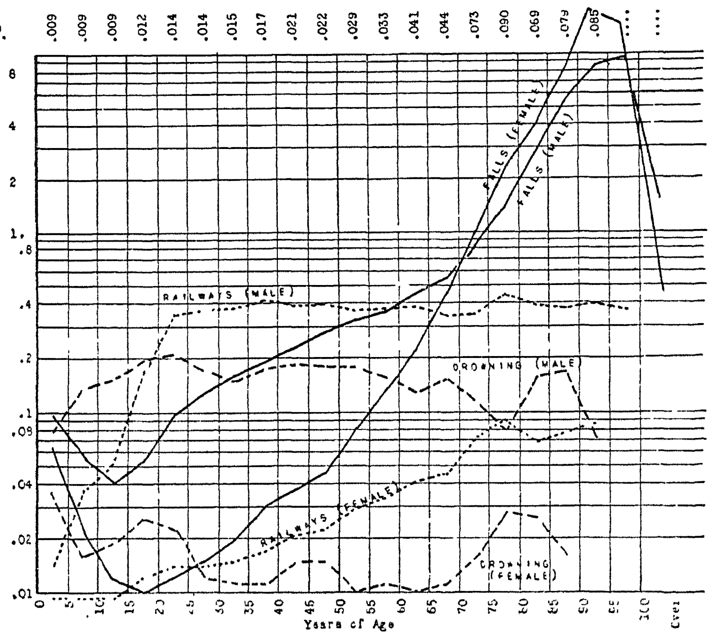


Fig. 343.

decrease" scale, though possible, is really little used and in actual practice more or less exceptional. In common with the examples given in the previous paragraph, it is always a little puzzling because the uniformity of "decks" has apparently been destroyed.

To the fact that the logarithmic charts cannot show a true zero point, we owe one of its most important and unique features, namely that the height of a curve upon this paper is entirely without significance and the curve may be moved

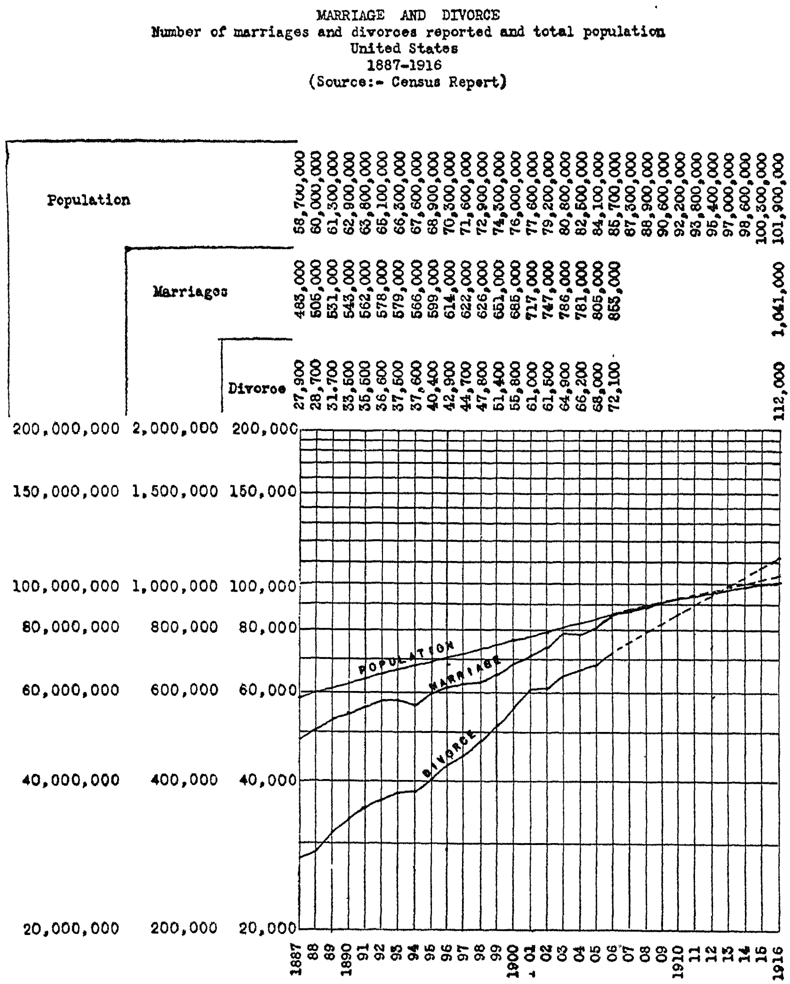


Fig. 344. Several Scales in a Single Split-deck.

bodily up or down upon the chart without altering the significance of the curve-fluctuations. When we say that the chart cannot show a zero point, we mean that you could continue the deck and ruling of the chart paper downward infinitely far without ever succeeding in reaching zero. You would merely reach smaller and smaller fractions of positive values. The true zero-point is located out in infinity. It is therefore taking no liberties with the chart to slide one curve up or down as far as you please to make it more easily comparable with another curve, because you are not really changing the position of the zero point (that still lies for both curves out in infinity). If the two curves are upon separate sheets of paper we may slide one piece further down than the other so as to bring the curves into close association with each other. Likewise if we plot both curves upon the same chart we can use a small scale for the lower curve and a much larger scale for the higher curve, and so superimpose one upon the other. (In this case separate scale figures may be an advantage to the reader but they are not essential.) This juxtaposition of curves is one of the chief advantages of the rate-of-change paper and can easily be carried so far as to make one curve cross or intersect the other. It is however considered better practice to prevent the crossing of two curves which have been brought together in this way, by sliding one curve lower down and altering the scale correspondingly. If you are intent upon a very clear exposition of the artificial nature of this juxtaposition, you can wipe out a small portion of the co-ordinates of the paper between the two curves so as to indicate a break or omitted portion of the chart-field between them.

From this arises a very important use of rate-of-change charts in the detection of correlation. For not only is the distance or interval between points upon the logarithmically projected chart useful for comparing the successive items in the same series of data, but it is also useful in comparing corresponding points upon different series. The problem here is not to scrutinize the slope of parts of one curve, but to scrutinize the distance between two curves. This distance is, as you remember, when a distinct parallelism or mirroring of the two curves is noticeable, an evidence of that similarity of behavior which is called correlation. To some extent, correlation can be discovered by the use of index numbers, or relative percentages, which make the fluctuation in different

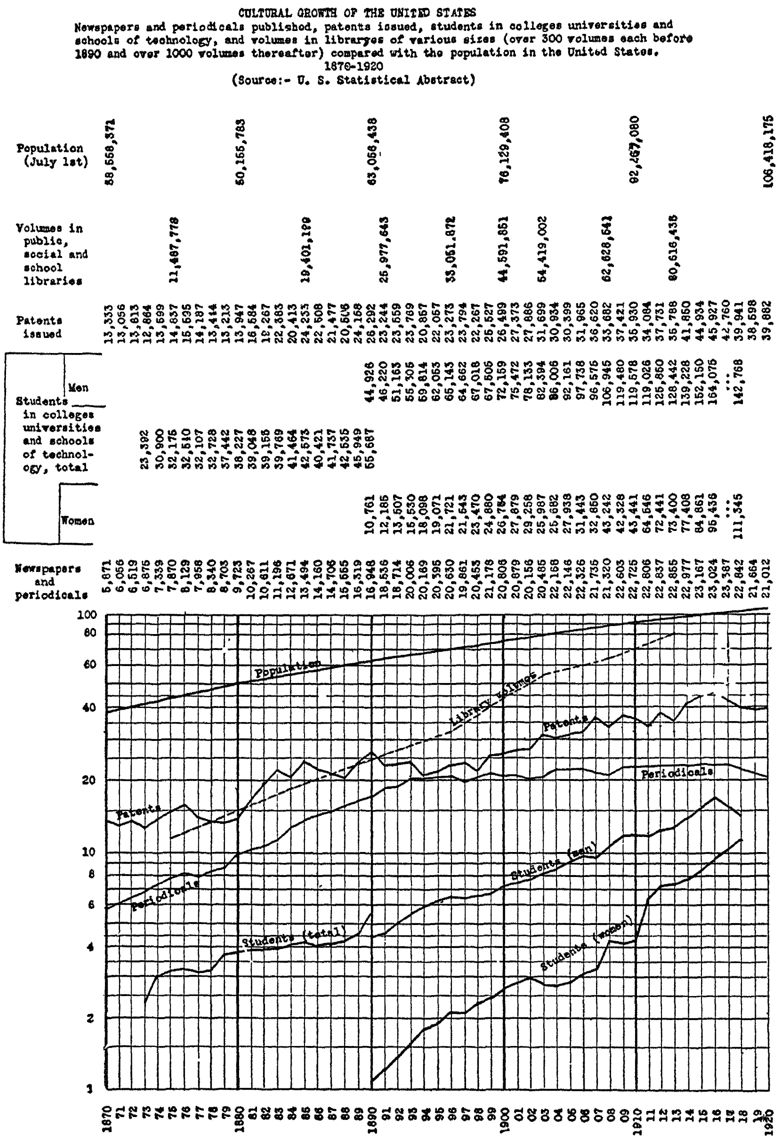


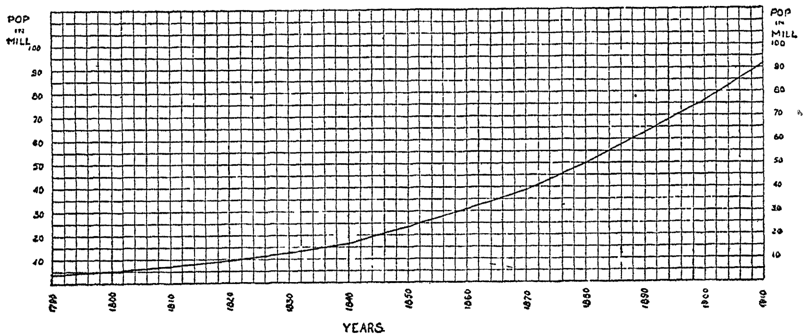
Fig. 345. Shifting Curves to Avoid Insignificant Crossings.

sets of data more comparable. But a far more precise comparison is afforded by the rate-of-change curve. To obtain an exact measure of the degree of correlation between various data, it is of course necessary to fall back upon the mathe-

matical operations which yield a "correlation coefficient." But for most purposes the graphic method afforded by rate-of-change charts is sufficient and it is always, of course, much easier and more rapid.

When a great number of series are plotted upon rate-of-change curves they can be compared in short order merely by inspection. If they have been plotted upon paper which is sufficiently translucent, closer inspection can be made by the use of "light analysis," that is by laying one chart over the other, holding the two of them up to the light, and sliding one back and forth and up and down until it most nearly coincides with the other. Mirroring can be detected by turning one of the charts upside down and bringing them together for light analysis. The great advantage of the use of the rate-of-change chart for correlation detection is due to the fact that various "lags" in the correlation of the fluctuation can be immediately corrected by this method, whereas by the long mathematical process, a slight lag might be sufficient to wipe out and conceal any correlation which might exist, even though that correlation be most complete. Even when the mathematical processes are to be used, in order to measure the correlation exactly, it is best to use the graphic method first, in order that the mathematical work may be performed only upon the series showing appreciable correlation and in order that any lag which may be present can be corrected for. In short, for the work of correlation studies, so essential to forecasting, the rate-of-change curve chart is becoming recognized as necessary. Extrapolation has already been discussed as a means of forecasting or predicting the nature of future developments and this paper will be found admirably adapted to such work, being statistically indicated wherever the fluctuations of data are logarithmically more regular than they are arithmetically.

A word may be said as to terminology. The curve charts which we have previously examined under the general name of "amount-of-change charts" are sometimes called increment charts or difference charts from the fact that the fluctuation of the curve plotted upon them represents increments added to or differences subtracted from their previous values. They are also sometimes called arithmetical charts, from the fact that the straight line upon them represents an arithmetical series. The charts to which we are coming and to which we



ACTUAL POPULATION OF THE UNITED STATES. DIFFERENCE METHOD.
Showing the impossibility of correctly comparing rates of increase at different periods.

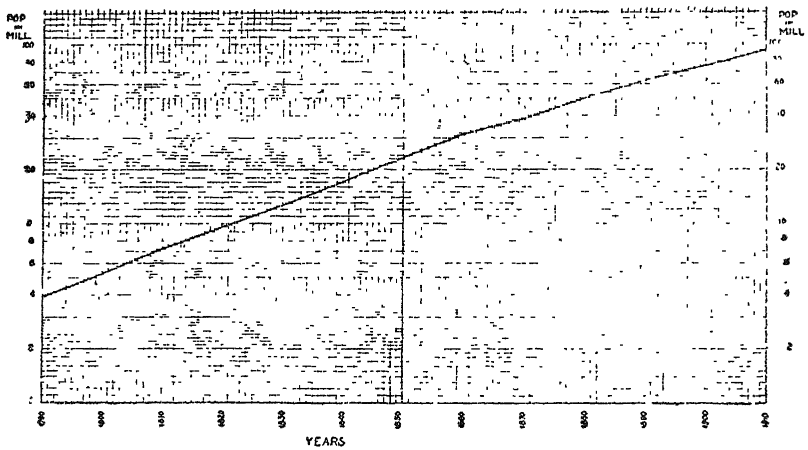
From Irving Fisher's "The Ratio Chart," in American Statistical Association Quarterly, June, 1917.

Fig. 346. An Amount-of-change Chart.

give the general name of "rate-of-change" charts are sometimes called ratio charts[1] from the fact that the fluctuation of the curve plotted upon them represents in a certain sense the ratios of change rather than the amounts of change. The name is a poor one because the ordinary bar-chart and the ordinary amount-of-change curve both express more graphically the actual ratios between the quantities; the rate-of-change chart showing graphically only the changes of ratios but not the actual ratios themselves. These charts are most frequently called logarithmic charts from the fact that they show logarithms and not the anti-logarithms or natural numbers.

The real distinction between the amount-of-change curve and the rate-of-change curve is the distinction between quantitative and qualitative analysis of data. For a quantitative analysis, that is a study of the actual quantities involved, the amount-of-change paper is necessary. But for a qualitative analysis of the figures, that is a study of their comparative relations, ratios, and proportions, the rate-of-change paper is necessary. The fundamental distinction to keep in mind is that this qualitative rate-of-change paper illustrates relative or proportional changes and is significant only as to them. Moreover, because this paper does not illustrate totals, it is always well to have at least your important data plotted both upon amount-of-change curves and rate-of-change curves, that from each type of curve you may easily get its particular sig-

[1] The name, it is believed, was introduced by Professor Irving Fisher, *obit. cit.*



THE SAME. RATIO METHOD.
Showing clearly the slight deviations, since 1860, from a uniform rate of growth.

From Irving Fisher's "The Ratio Chart," in American Statistical Association Quarterly, June, 1917.

Fig. 347. A Rate-of-change Chart.

nificance. And it is because the qualitative study, that is the study of relative values, is so frequently the more useful one that the rate-of-change charts are themselves so commonly more valuable.[2]

[2] The significance of the logarithmic projection need not be difficult to understand. The measurement of star-light in magnitudes, and the measurement of sound-waves by means of the octave and its parts, are familiar examples in which we have adopted logarithmic (or exponential) units of measurement, as clearly necessitated by the type of the phenomena. It is permissible to think, therefore, that in all cases where the logarithmic projection is found suitable, nature is operating in logarithmic units, that is, changing organically in geometric progression, while man is still thinking in terms of arithmetical units, and must needs project these logarithmically.

Chapter XXXVI

HISTORICAL RATE-OF-CHANGE CURVES

For the curves of historical data we find a type of rate-of-change chart which is fast increasing in popularity and general use. Though it involves the logarithmic projection of scale, yet that fact is so completely camouflaged by the rulings of the chart (in "decks" as described in the last chapter) that we need no longer apologize for its use in a popular publication nor attempt to explain it in conferences. We merely murmur something about its being a truer picture of fluctuations and let it go at that. If the other chap does not understand—and this includes chief executives and officials—he at least realizes that he ought to understand and enters no protest. In short, this chart form has already reached the stage of notoriety in which it need no longer skulk about in laboratory corners, but can parade in public with a slightly exclusive, but very effective manner.

The peculiarity of this chart is that it has a logarithmically projected scale along one axis only, the vertical or y-axis. Its x-axis is innocent as a new born babe of any such development—that is to say, its x-axis is projected arithmetically. And for this reason the chart is often, in charting office parlance, somewhat crudely but tersely called "semi-logarithmic." Technically, of course, it can only be described as "x-arithmetic, y-logarithmic." That the form is appropriate, however, for historical data, may be seen if we recall the law of organic growth in which it was stated that growth by uniform percentages took place at uniform intervals or periods of time. Indeed, as the law of organic growth essentially deals with historical data, that is, data during various points or periods of time, we may consider that chapter a discussion of the general theory of the historical rate-of-change chart in particular.

The student may be surprised that time, in itself a natural phenomenon, should defy the law of organic growth. Eminent

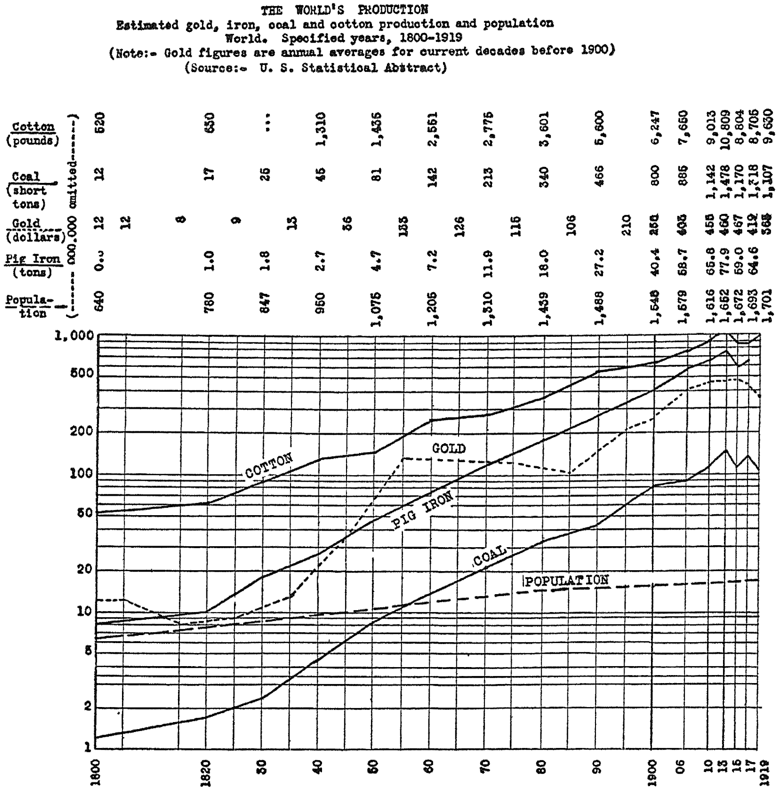


Fig. 348. Long-Time Series of Economic Data.

engineers with exceptional mathematical ability, have questioned the logic of charting time arithmetically with logarithmically projected dependent variables. The answer may be found in the very close approximation of the logarithmic series to the arithmetical one through small ranges. In the whole history of time the origin is so infinitely far removed and the range of known history so extremely minute (comparatively speaking) that there would be no appreciable difference in the resulting chart were the x-scale plotted by either method. While it is certain that for modern times the two would coincide, it is furthermore impossible to fix the true origin of time. For zoological and palaeontological charts the birth of the moon, or some similar event, might be a useful zero-point, but in business and in historical charts in general the result would still be the same. In other words, we may consider the his-

VIOLENT-DEATH RATES
United States, 1900-1920
(Sources:- For Homicides, suicides, total accidents and automobiles, the Census Reports;
For Lynchings, the Tuskeegee Institute; for street accidents, Dept of Health, N. Y. City.)
(Note:- Figures in parenthesis are annual averages for 1901-1905.)

Street accidents New York City	0.14	0.17	0.24	0.32	0.36	0.89	1.53	0.97	1.81	1.81	2.32	2.63	3.79	5.80	6.03	6.62	7.66	9.72	12.35	13.76	13.47
Automobiles																5.8	7.3	8.9	9.1	9.4	10.4
Railroads	10.0	10.9	10.8	12.2	10.2	11.5	12.4	13.6	11.5	9.6	10.5	11.1	11.1	11.3	10.5	8.6	9.2	9.9	9.0	6.7	6.6
Total accidents and unknown	79.0		(84.9)						81.2	80.2	84.3	84.6	82.4	85.3	78.5	76.3	83.9	87.2	80.5		
Suicides	11.5		(13.9)						17.8	16.5	16.0	16.2	16.0	15.8	16.6	16.7	14.2	13.2	11.8		
Homicides	2.1		(2.9)						6.4	5.6	5.9	6.6	6.5	7.2	7.3	6.9	7.1	7.6	6.7		
Lynchings	.137	.174	.121	.126	.106	.078	.078	.072	.097	.077	.080	.076	.067	.054	.053	.069	.053	.037	.061	.078	.067

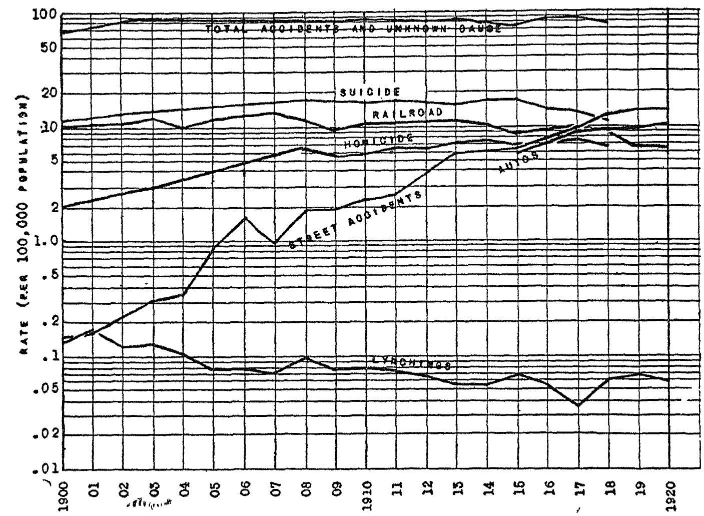


Fig. 349. Short-Time Series of Economic Data.

torical rate-of-change chart to be really logarithmically projected on both axes, with the x-axis scale range, however, so short that it appears to be arithmetical.

In the construction of historical rate-of-change charts, the same principles apply (with the exception of the logarithmic scale on the y-axis) as in the construction of historical amount-of-change charts, to which several chapters have been devoted. A standardized form is useful when many charts are being made, most of the marketed forms described in the last chapter being useful for this purpose, as they contain no ordinates or vertical rulings and can easily be provided with the latter

to suit each case. The field should not occupy more than half of the page (on an $8\frac{1}{2}$ by 11 basis) in order that data can be entered in full. The same position of the field in the lower right hand corner should be maintained, when the charts fit together to form a series, in order that they may be closely overlapped either horizontally (to show one continuous curve) or vertically (to show seasonal and cyclic fluctuations). The same problems as to plotting points of data at unequal time intervals arise and they must be settled in the same way. In general the historical rate-of-change chart is but a duplicate of the historical amount-of-change chart, modified in regard to its y-axis scale and the plotting of its dependent variable data.

Not all the historical amount-of-change charts, however, can be duplicated in this way. With the simple curve, there is of course no difficulty (either in its usual or silhouette form). But in the Zee-chart the cumulative is not profitably copied; indeed it is rarely of any use to plot historical cumulatives (at least when these are for limited and repetitive cycles or periods of time) logarithmically. The bar-charts (except silhouette bar forms of curves) should never be plotted logarithmically, as by their nature they suggest and imply a significance in their heights, and the logarithmic chart, as you know, is without significance in the height of its curves.[1]

[1] It is here perhaps best to warn the student against too confident use of other chart methods for the purpose of showing rate-of-change fluctuations. In previous chapters the series of chain-relatives or chain-percentages has been discussed and it has been pointed out that the chain relatives are the anti-logarithms of the logarithmic differences between successive items in a historical series. As the rate-of-change chart shows by the fluctuations of its curve the amount of these logarithmic differences, it is easy to see that the fluctuations of a rate-of-change curve do not represent the chain-relatives or chain-percentages, but represent the logarithms of these chain-percentages. Attempts have therefore sometimes been made to present graphically the chain-relative series itself. A little experimentation however will show that the method is not for most purposes useful. To be strictly accurate, the chain-relative series must not be shown by a single continuous curve but by a number of successive curves in which each plotted point is connected with the previous ordinate at its intersection with the 100% line. The result is a wholly disconnected picture and is to some extent liable to wholly meaningless changes of form when the time units of the data are shifted or changed. Nor is it possible to present a more connected story by joining each of these disconnected curves and plotting these new points the proper distance above or below the last preceding points instead of above or below the 100% line, for this would amount to a cumulation of the chain-relative series, additively, when the series properly cumulates only by multiplication, and it would result in a curve in which no uniform scale is possible for the dependent variable, and in which points at the same height upon the paper have different meanings or values at different portions of the series. Except in very special circumstances the rate-of-change curve chart is the proper method for showing the rate-of-change of fluctuation in data.

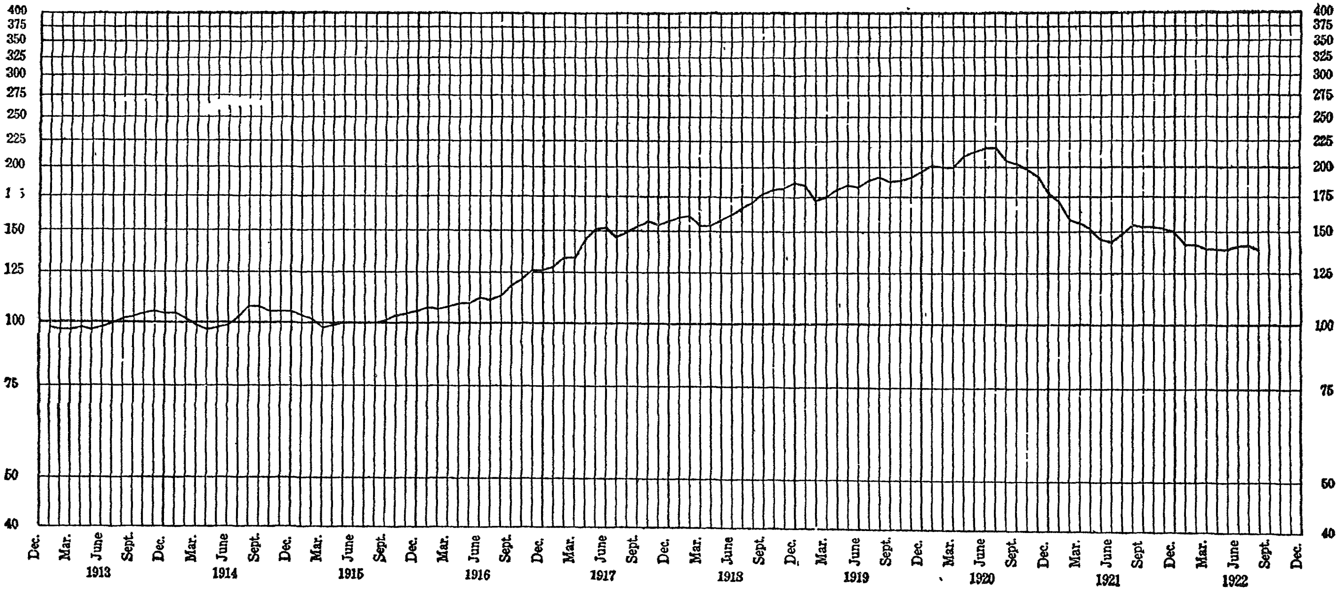


Fig. 350. A Careful Positioning of the Curve.

Showing retail price of all articles of food, combined, in the United States.

It would have been even better to begin the chart with 50 for the base line, as the 200-line would then be just twice the height of the 100-line and the similarity to an amount-of-change curve would have been more complete.—*From the Bureau of Labor Statistics "Monthly Labor Review."*

Where the charts are to be used popularly it is well to go to considerable pains to lessen the misunderstanding of the chart by people who do not understand its logarithmic nature. Some misunderstanding is bound to arise, but one of the simplest and most effective measures is to position the curve (where one only is shown) at about the height upon the paper at which an amount-of-change curve with similar fluctuations, would appear. Thus if your curve covers a range on the y-axis of let us say from the 100% point to the 200% point, the highest point on the curve is obviously double the absolute value of the lowest. Now if we cut off our chart-field at the 50% point, then clearly we will have equal distances between the bottom of the chart, the low point on the curve, and the high points on the curve. Now when Mr. Average Man—and it may be Mr. Average Congressman and seem very important to you—picks up this chart, he goes through the following motions: "Ah, one of those damned curves—what the deuce does it mean? Oh, profits on rotten meat. Well, I see they doubled during the war. I can understand this thing easily." And then he drops it again. He does not notice that the bottom of the chart registered 50% and not zero on its vertical scale, but that does him no harm. He has at least seen one thing truly, the relation between high and low points.

In business reports it is always a good plan to present both rate-of-change and amount-of-change curves simultaneously for all important historical data. The two charts of the same information should be face to face so that the reader is confronted with both at once, and cannot mistake the rate-of-change form or judge quantities by its curve. Nor is this useful merely to avoid mistakes on the part of readers unconversant with rate-of-change paper. There is often a real use for quantitative analysis of the data, which can be seen only from the amount-of-change paper. The rate-of-change chart is indeed for most purposes more effective, but it fails wholly to give any picture of total sizes or quantities involved, and is not the panacea which some of its enthusiasts would have us believe. The whole truth about a series of data requires not only the rate-of-change but also the amount-of-change method.

In more scientific reports and in records which will only be used by those who cannot misunderstand them it is often useful to combine the two types of curve for the same data

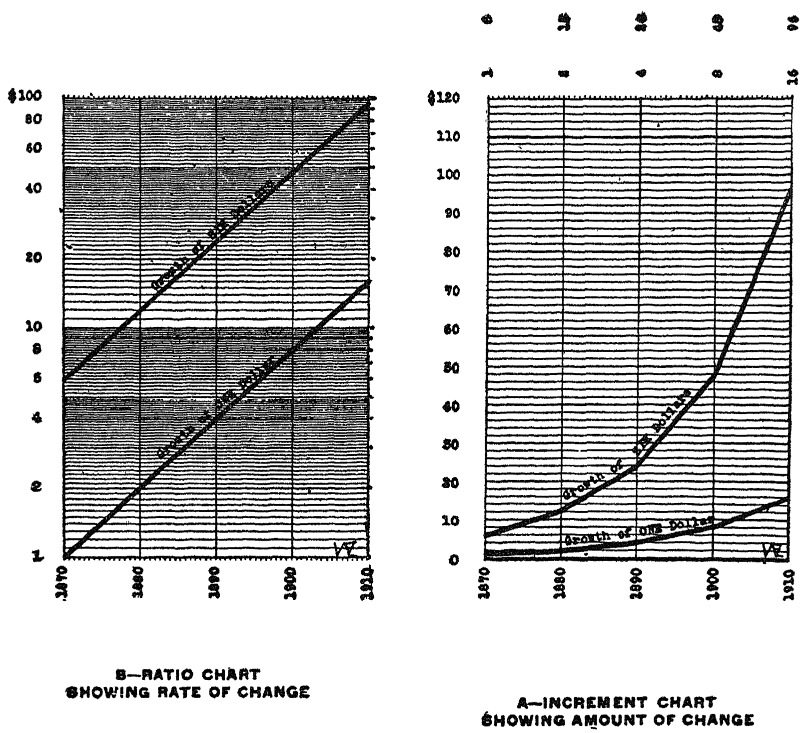


Fig. 351. Comparison of Rate-of-Change and Amount-of-Change Curves. The chart shows the growth of one dollar (lower curve) and six dollars (upper curve) at compound interest.—*Permission of Mr. John Wenzel.*

upon a single chart, superimposing one upon the other. The simplest method of doing this is to plot one curve for the natural numbers and another curve upon the same chart, with the same arithmetical rulings, for the logarithms of the data. The first or natural number curve can be labelled *Y* and the second or logarithmic curve can be labelled Log *Y*. If the paper used for the chart is fairly translucent, it is not necessary to look up the logarithms in plotting this second curve, but a sheet with heavy logarithmic ruling can be placed underneath the chart and the two sheets held up together to the light while the logarithmic curve is plotted, from the logarithmic ruling on the lower sheet, which are not intended to appear upon the chart itself. The slide-rule will serve equally well, if its scale is appropriate, an engineer's scale being used for the arithmetic curve and the slide-rule for the plotting of the logarithmic curve. A clearer method of show-

ing the distinct nature of these two superimposed or combined curves than by merely labelling them Y and Log Y, is to trace or draw in a small portion or zone of typical logarithmic ruling immediately around the logarithmic curve, thus showing that the latter has been cut out from a logarithmically projected chart and inserted upon the arithmetical chart. Needless to say, the combination of these two wholly different types of plotting of a curve should be made upon a single chart only when that chart will show a single series of data, for obviously several curves brought together in this way would become confusing.

No discussion of the historical rate-of-change curve is complete without mention of its value in forecasting. The use of

TRADE UNION MEMBERSHIP OF THE WORLD

Number of members in 20 countries
1910-1919
(Source:- International Labor Office)

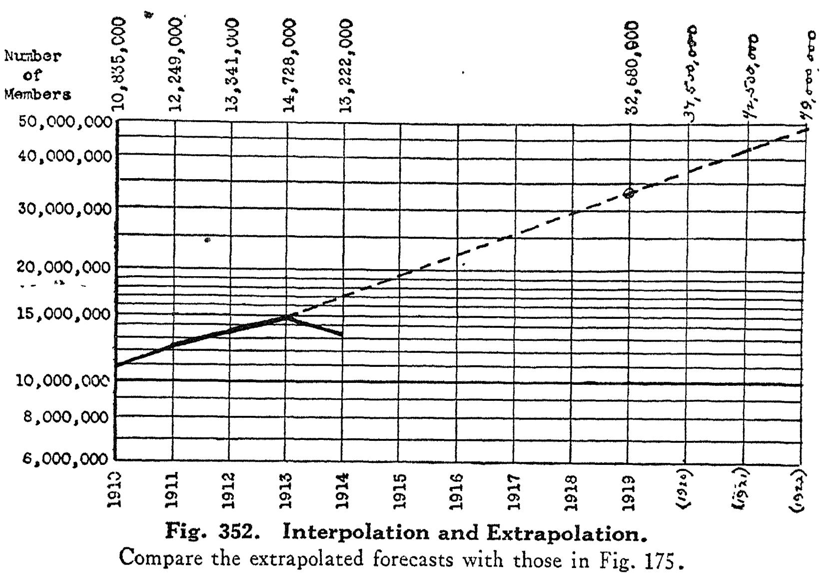
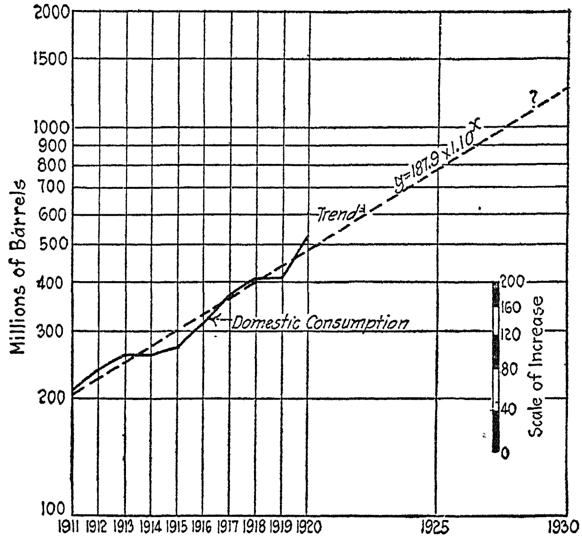


Fig. 352. Interpolation and Extrapolation.
Compare the extrapolated forecasts with those in Fig. 175.

the logarithmic projection in general for correlation studies has already been noted in the last chapter. But there is a very special usefulness of the rate-of-change chart in predicting from the course of past events, the probable course of future events. The curve is merely extrapolated to the ordinate of

the desired point in the future and the intersection point is read from this extrapolation. The degree of success or reliability achieved in such methods of prediction depend of course upon the faithfulness with which the future developments follow the course of the past record. But an estimate of the probability of this can often be made from an inspection of the past course of the curve itself. The method is only statistically indicated where the curve of the past shows very regular and uniform trend. Through a combination of causes, it often happens that these estimates can be more reliably and usefully made by extrapolation of the rate-of-change curve



From Joseph E. Pogue's "Economics of Petroleum."

Fig. 353. Careful Extrapolation.

A straight line (dotted) has been fitted to the curve and projected ahead ten years to give forecasts. In 1930 the forecast is about 1250 million barrels as against 800 millions indicated by extrapolation on amount-of-change paper (see Fig. 176).

than by extrapolation of the amount-of-change curve, for not only is the trend of the past often more uniform upon the former, but also in general it is true that the phenomena which show uniform rate-of-change in the past can be more relied upon to maintain their trend than the phenomena which show uniform amount-of-change.

The historical rate-of-change curve is indeed one of the most important instruments in the treatment of ordinary

business statistics and economic studies. It has not yet reached the height of its vogue but it is fast gaining in prestige and is fully deserving. You will not only find it illuminating and fascinating to work with, but you will soon discover, if you have not previously used it, that it opens up to you a new world of investigation and research.

CHAPTER XXXVII

LOGARITHMIC FREQUENCY CURVES

Hamlet: Do you see yonder cloud that's almost in shape of a camel?
Polonius: By the mass, and 'tis like a camel, indeed.
Hamlet: Methinks it is like a weasel.
Polonius: It is backed like a weasel.
Hamlet: Or like a whale?
Polonius: Very like a whale.

This same doubt and debate arises over the shape of every frequency curve. Such curves commonly present the widest variety of shape and contour. It is indeed possible to conceive of frequency data for almost any curve which may be imagined. Some attempts have been made to classify these various forms and divide them into a few typical groups. These classifications rest upon the relative positions (along the x-axis) and magnitude (or amplitude) of the peaks and valleys in the curve. They are restricted to simple curves, presenting not more than one peak and two valleys (strictly speaking, half-valleys) or one valley and two peaks (half-peaks). Where more peaks or valleys occur in the curve, and these cannot be smoothed out by applying a process similar to the moving-average or total process, it is rather assumed that the curve is not simple, but compound, being really a combination of two or more separate simple curves. The breaking down of a composite curve into simple curves is a problem of advanced and difficult mathematical steps, superficially not unlike harmonic analysis, and cannot be discussed here. Outside of the engineering and scientific fields, these curves will be rarely met and the student of business and economic statistics need only acquaint himself with the treatment of simple frequency curves. We shall first consider these simple frequency curves as they appear on amount-of-change plotting paper.

While no classifications have proved very useful, the best which has so far been formulated, and one of the simplest, is the classification of Professor Yule.[1] Yule finds four common types,

VITAL SUPERIORITY OF THE FEMALE
Percentage Excess of Male over Female Death Rate
England and Wales
1851-60 and 1901-10
(Source:- Reports of Registrar General of England and Wales)

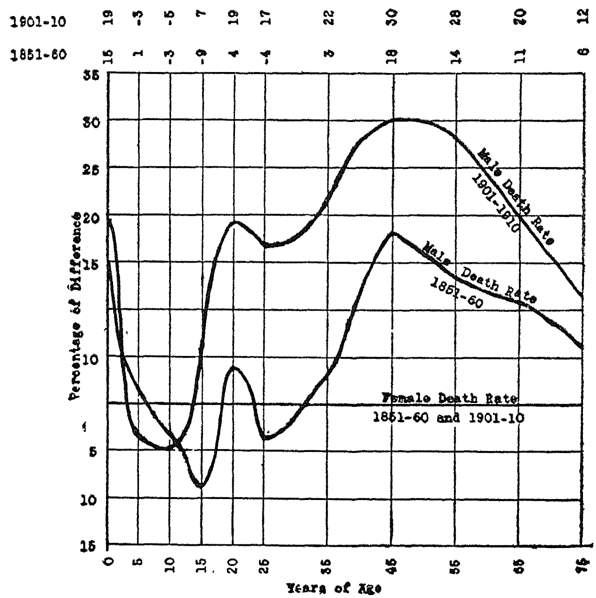


Fig. 354. Compound Curves.

which he calls, respectively, the symmetrical curve, the moderately asymmetrical curve, the extremely asymmetrical or *J*-shaped curve, and the *U*-shaped curve. These names are self-explanatory. In the first or symmetrical curve, there is a single peak, from which on both sides the curve slopes away symmetrically. The salient facts about this curve are its width or spread ("range"), the height of its peak ordinate, and the relative heights of its ordinates at regular or certain irregular ("quartiles," "decils", or "percentiles") intervals away from the peak ordinate. The last detail would tell us whether the peak was narrow, indicating great concentration of the observations about that point, or wide, indicating a scattering or "disper-

[1] *Cf.* Yule, *Theory of Statistics.*

sion" about the point, and there are mathematical methods of expressing these details which may be found by consulting the statistical authorities.

OUTPUT OF COAL MINERS
Average Number of Tons of Bituminous Coal Mined by Pick-miners per 8-hour day in 118 mines, 17 states (all fields)
United States
1919
(Source:- Ethelbert Stewart)

Number of Tons	Number of Miners
Under 3	577
3 - 4	913
4 - 5	1251
5 - 6	1491
6 - 7	1395
7 - 8	1275
8 - 9	984
9 - 10	672
10 - 11	462
11 - 12	290
12 - 13	171
13 and over	336

Fig. 355. A Frequency Series Which Appears Slightly Asymmetrical.

The second, or moderately asymmetrical curve, presents the same general form as the first, but its peak is no longer in the middle of the curve, being nearer to one end than the other. Consequently the curve does not fall away on both sides with symmetry. Here we have a new element which is statistically known as "skewness", the curve being skewed over to one side, and there are mathematical methods of describing more or less explicitly the degree of this skewness.

In the third or extremely asymmetrical type of curve, we generally have but half a peak and half a valley, that is, the curve begins at a peak on one side and slopes down toward a valley, but the other half of the peak, the opposite slope, as it were, may be lacking. Where the opposite slope is present, it is very near to a straight vertical line and the curve is merely extremely asymmetrical. When it is lacking, the curve is called *J*-shaped and not only the lower or valley end, but also the peak end, of the curve may be "asymptote," to the axis of the chart. By an "asymptote" is meant a line which, while

approaching infinitely near to an axis, gives no promise of actually meeting it, no matter how far it be extended outward along the axis.

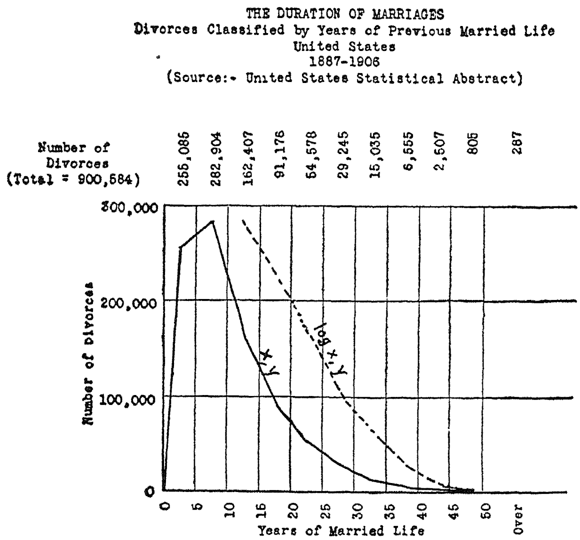


Fig. 356. An Asymmetrical Distribution.

The dotted line is the same curve plotted upon a logarithmic horizontal axis. Note that, if this plot could be extended through the first open group, it would probably become approximately symmetrical.

The fourth type, the *U*-shaped curve, is one which describes a letter *U*, being made up of a valley and two half-peaks. And though it may be a very sharp or pointed, V-shaped valley, it is perhaps more frequently found to be more or less flat, just the reverse being generally true of the other three types. And it might be added that this *U*-shaped curve may be found in both symmetrical and asymmetrical forms, according to the presence or absence of skewness. Yule, however, makes no subdivision of the U-shaped curves, as they are comparatively rare.

It is now our purpose to show that these four more or less distinct types are often interchangeable forms, which can be evolved by different statistical and graphical processes, from the same original series of data. For this purpose we naturally regard the symmetrical type of curve as the more desirable form, and shall consistently refer other forms back to it. The reason for this, as has already been indicated, is that the more

regular and symmetrical a curve be, the safer and more trustworthy would seem its use for interpolation and generalization. The symmetrical form is not significant only to the mathema-

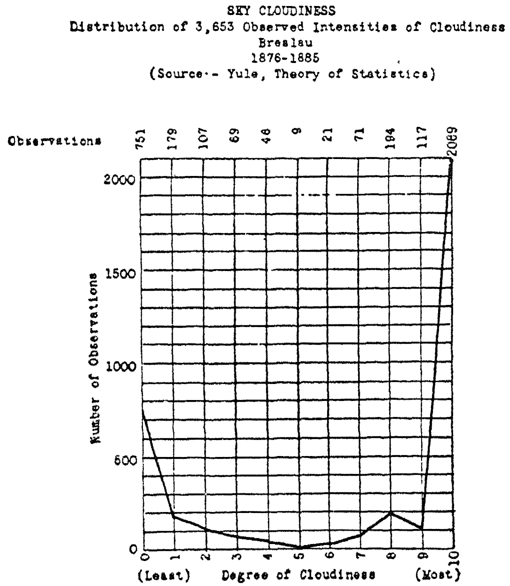


Fig. 357. Yule's Example of a U-shaped Distribution.

tician who will seek a general equation for the phenomenon described on the chart, but it is also more significant to the layman, who, having once seen the symmetrical curve, will not so easily forget it. And at this point we enter the subject of the logarithmic projection of the scales.

To begin with the second, or moderately asymmetrical form of frequency distribution, we have to note that its curve can often be made symmetrical by plotting upon a logarithmic scale. Plotted on rectilinear co-ordinates, that is, on the amount-of-change plotting paper, the right-hand "tail" of the curve, that is the slope toward the half-valley at the right-hand side, is generally longer than the one at the left. At the same time, the values of the independent variable, that is, the values of the points along the x-axis, are larger at the right-hand end of the scale than at the left. So it is obvious that a logarithmic projection will shorten this part of the scale in comparison with the rest of the scale (for the log projection always condenses the larger values). The result is to shorten the longer tail, often

enough to produce absolute symmetry in the two slopes of the curve.[2] Nor does this process need to be wholly empirical, for the student will soon learn to detect in advance the curves

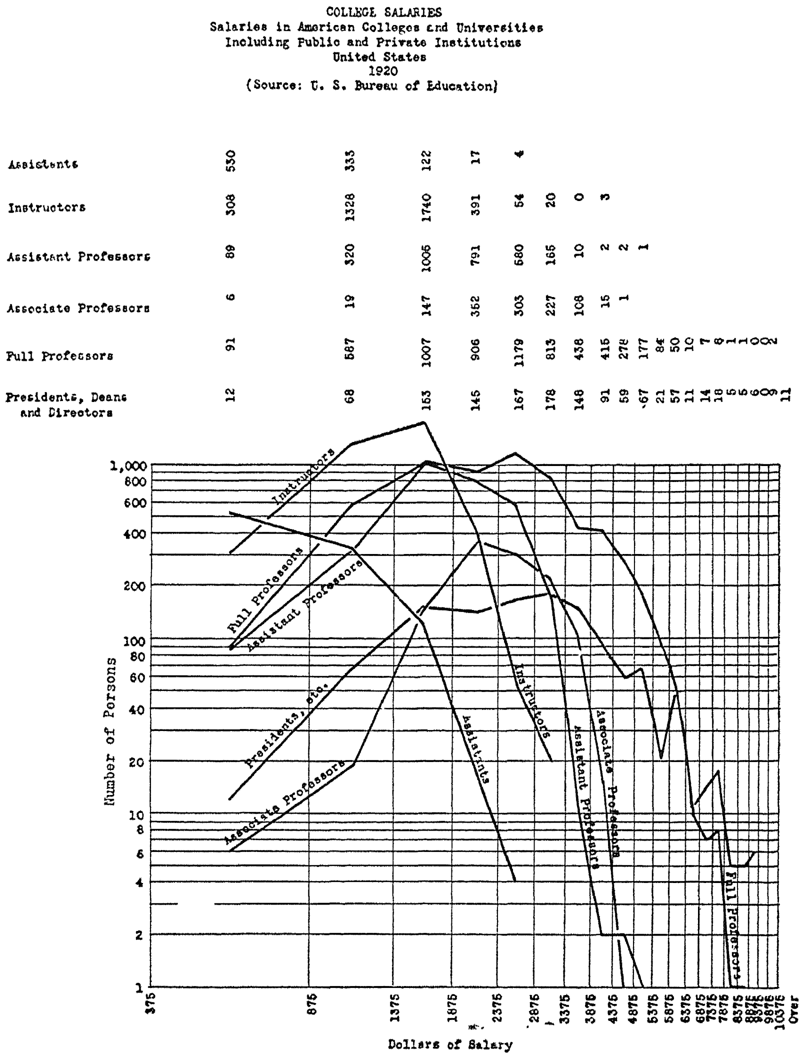


Fig. 358. Six Moderately Asymmetrical Distributions.

Note that these have been made more symmetrical as to their sides and more rounded as to their peaks by means of logarithmic scales.

---

[2] See Fig. 361 on p. 435.

which, by this treatment, may become symmetrical.[3] The success of the method depends, of course, upon the nature of the independent variable, whose values appear as calibrations or scale-figures along the x-axis of the chart.

From the above it is clear that it would be impossible to apply this method directly to data in which the independent variable includes both negative and positive values and crosses through zero. When zero is not a limit of the range, but is included inside the range, it would be necessary, if the method is to be applied, to change the variable to values which are wholly positive or negative. This can easily be done, if the zero-value proves on inspection to be purely nominal, arbitrary, or relative, and to have no real zero meaning. Thus if the series

RENT INCREASES
Number of Families Reporting Increases of Rent
Government Employees, Washington, D. C.
Year Ending Oct. 1, 1920
(Source:- Monthly Labor Review)

0	1,983
1 - 9	83
10 - 24	648
25 - 49	480
50 - 74	142
75 - 99	41
100 and over	29

Fig. 359. The Independent Variable is Measured from an Arbitrary Zero Point.

be so arranged as to show deviation from its mid-point (median), mode, mean, or from some other particular value, the deviations, being measured above or below this point, will show in the table as positive and negative stubs, and the point itself will show as zero. By merely adding to all stubs the true value of this point, we return the data to its primary form and wipe out the false zero-value. In other cases the zero-value, though arbitrary, cannot be so easily given its true value, and the work is more difficult. When zero is a limit of the range of the data, and is actually met in some of the data, the log-chart can be used, but of course the values of the curve in the first interval (next the zero point), cannot be plotted, as that interval becomes

[3] See Chapters XXVII and XXVIII.

infinitely long.[4] In the same way, the final class or group cannot be plotted where it is indeterminate, for it too is infinitely long.

In general, the same considerations apply to the possibility and usefulness of the logarithmic projection of the x-axis scale as are applied to the logarithmic projection of the y-axis scale, already discussed in the foregoing chapters. The logarithmic projection would seem appropriate whenever zero is an infinitesimal limit (approached but never reached) to the independent variable. Such a condition is inherent in the nature of the phenomenon and can be detected immediately therefrom. Most economic data are susceptible to the process of log projection. It is noteworthy that each of Yule's examples of the moderately asymmetrical curve can be plotted on logarithmic paper and made symmetrical thereby. Human beings, for example, cannot have a negative height, nor a zero height; communities cannot have zero populations; manufacturing establishments cannot have zero employees; nor farms, zero acreage. Frequency series of such phenomena, classified as to their sizes, seem to call for logarithmic projection *a priori*. On the other hand, profit or loss can be negative, as can net worth and balances of all sorts, including "stock" or "fund" data, time and space dimensions with reference to particular points, and for such data the log projection would ordinarily be both impossible and meaningless.

Whether or not the vertical or y-axis be given a log projection is relatively unimportant. Ordinarily when the x-axis is so plotted, the y-axis can be also, and the resulting curve shows a more rounded, less pointed peak. This may or may not be desirable. For identification with the normal curve described in the next chapter, perhaps more often the arithmetical projection of the y-axis, even with the log x, is desirable, but it may be that the identity will be established only when both y and x are logarithmic.

The third, that is, the extremely asymmetrical or *J*-shaped curve, presents two possibilities. If it is extremely asymmetrical, but has two "tails" or half-valleys about its peak, it may belong in the same category as the second or moderately asymmetrical curve. The difference in the degree of skew, which seems so much more violent in the third than in the

[4] See Fig. 356 on p. 429.

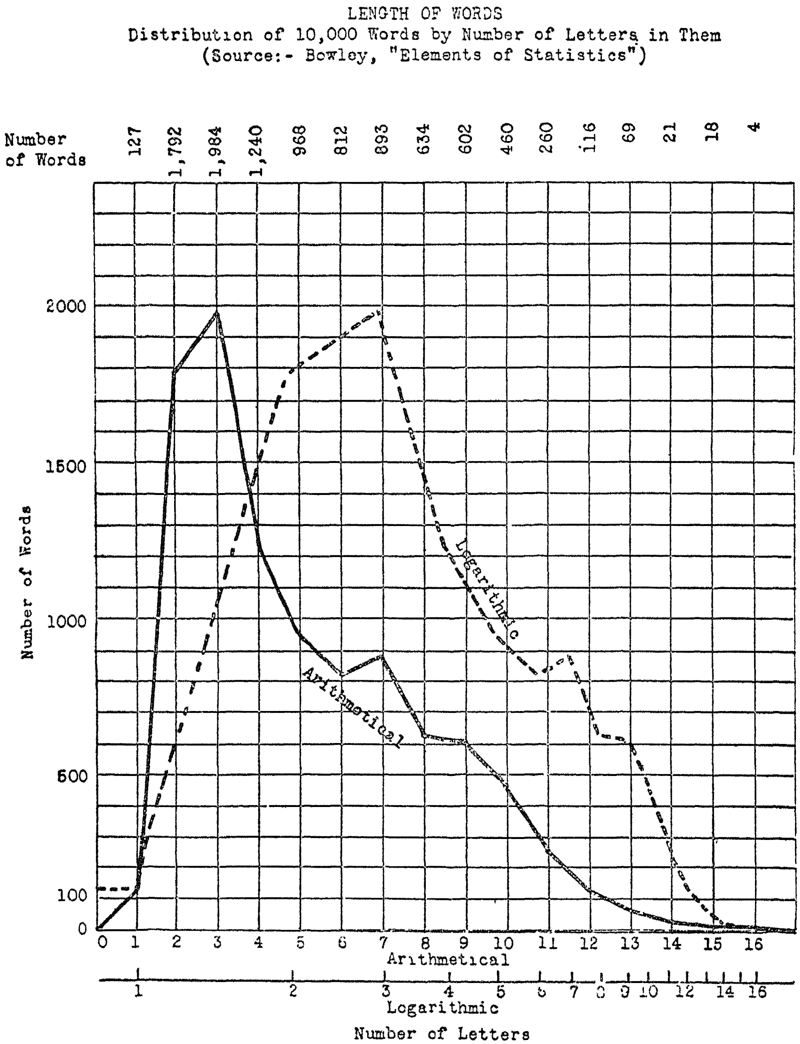


Fig. 360. A Moderately Asymmetrical Distribution which the Logarithmic Scale Has Not Made Entirely Symmetrical.

second, may be found wholly due to the greatly extended range. Thus if we classify farms by acreage, our table may include farms of less than three acres and farms of more than a thousand acres. Here the range is very great, being through three logarithmic decks, and the arithmetical projection has obviously brought the log mid-point very, very close to its left end (and

with it the peak, resulting in extreme asymmetry. The height of human beings varies through no such great range and hence shows, when arithmetically projected, only mild asymmetry.

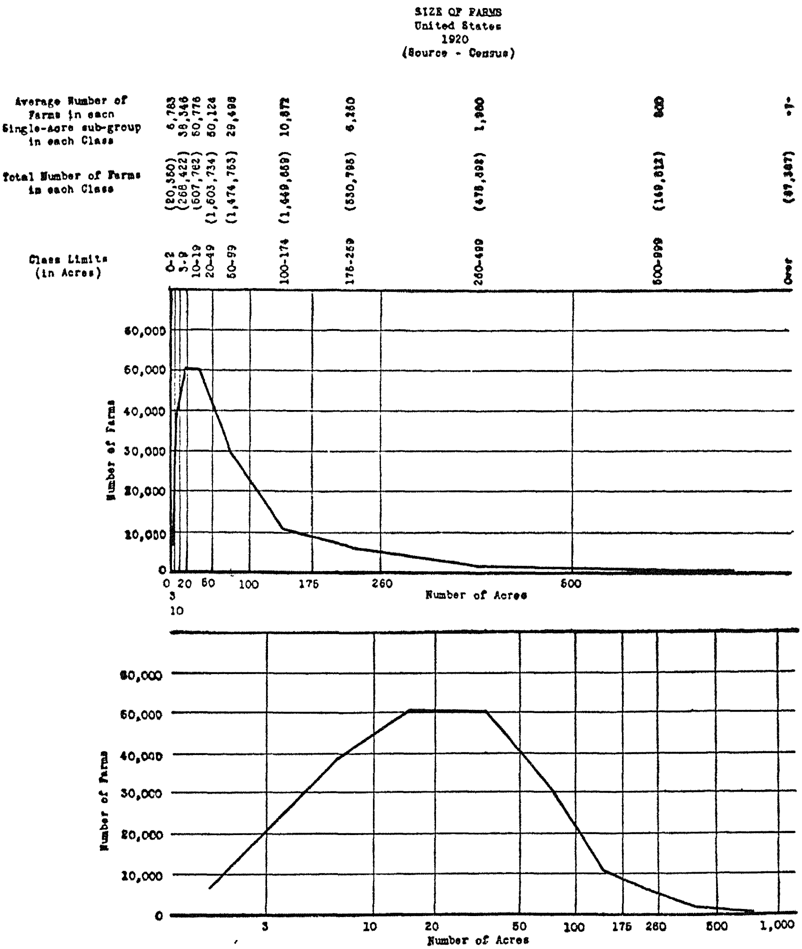


Fig. 361. An Extremely Asymmetrical Distribution Made Symmetrical by the Logarithmic Projection.

Note that the left "tail" of the lower curve has been extended through the open class (zero to two acres) as a straight line, and therefore obscures the symmetry.

Large ranges are extremely common in business and economic statistics, and the data will nearly always fly storm signals indicating the need for logarithmic projection. Thus the usual classification of cities by sizes runs through intervals with

SIZE OF STRIKES
Number of workers involved in strikes
United States
1916-1921
(Source:- Monthly Labor Review)

Classes: No. of workers		1-10	11-25	26-50	51-100	101-250	251-500	501-1000	over
No. of equiv. sub-classes		1	1.5	2.5	5	15	25	50	
1916	Class Total	197	345	412	413	395	348	238	254
	Class Average	(197)	(230)	(164)	(82.5)	(26.3)	(13.9)	(4.8)	
1917	Class Total	164	296	341	358	358	284	193	286
	Class Average	(164)	(197)	(136)	(71.5)	(23.9)	(11.3)	(3.86)	
1918	Class Total	143	268	334	344	371	278	141	216
	Class Average	(143)	(179)	(133)	(68.8)	(24.7)	(11.1)	(2.82)	
1919	Class Total	170	279	333	382	465	339	205	376
	Class Average	(170)	(186)	(133)	(76.4)	(31.0)	(13.5)	(4.1)	
1920	Class Total	150	299	313	326	341	262	136	196
	Class Average	(150)	(199)	(125)	(65.2)	(22.7)	(10.5)	(2.72)	
1921	Class Total	219	286	252	214	216	153	101	140
	Class Average	(219)	(191)	(101)	(42.8)	(14.4)	(6.13)	(2.02)	
Approx. number of workers		"6"	"18"	"38"	"75"	"125"	"375"	"750"	

Number of Strikes

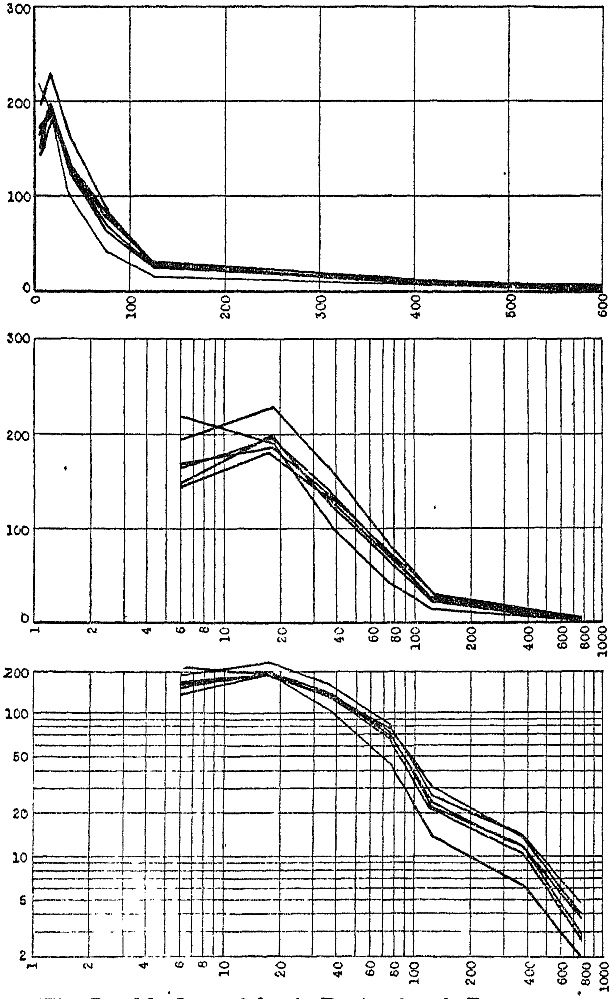


Fig. 362. The Double-Logarithmic Projection is Best.

various division points, such as 1000, 2500, 5000, 10,000, 25,000, 50,000, 100,000, 250,000, 500,000, 1,000,000. Who cannot see at once that these intervals are merely round numbers approximating, as closely as convenient, equal geometric, and not equal arithmetical, intervals?

When, however, the curve is *J*-shaped, that is, has only one tail, another possibility creeps in. For it may be that what we are treating as a *J*-shaped frequency distribution is really an ogive, that is, a cumulated frequency series. The cumulation of a distribution which is symmetrical on an arithmetical projection is easily detected, of course, but the cumulation of other forms may be neatly disguised in the description of the series and pass for a time unnoticed. The behavior of these other cumulatives we will discuss shortly, but it should be borne in mind that the most frequent example of the *J*-shaped curve is an ogive or curve of a cumulated series. And when you meet a *J*-shaped curve, examine it first to make sure that it is not an ogive. If it is not, then the possibility remains that it is a very extreme asymmetrical curve, which is really not *J*-shaped at all, and has two perfectly good tails, but one of them is so very short as to be swallowed up in the peak.[5] Thus Yule's illustration of the *J*-shaped curve, being the uncumulated distribution of personal incomes in Great Britain, is as he himself says, not really *J*-shaped, but merely so asymmetrical that its lower portion has been swallowed up in the mode of the series.

As to the fourth type, the *U*-shaped curve, rare as it is, it presents two or three obvious possibilities, one or another of which may serve in its analysis. In the first place, we must note that it may be merely an approximation to a distinct "yes or no" tabulation, its two terminal maxima representing the two alternatives and its intervening minimum the more infrequent compromises. Thus a tabulation of eyes by degrees of blue or brown color might show many wholly blue and many wholly brown eyes with relatively few eyes of the various intermediate shades, if the whole matter of color be a resultant of the presence or absence of dominant color determinants. Yule's illustration of this type of curve suggests a similar condition, being a record of sunshine and cloudiness at Breslau. If clouds be a local evidence of a falling barometer over a wide area,

---

[5] See Fig. 362, page 436, where this has happened in the last year.

and clear skies of a rising barometer, since it is obvious that the air pressure can shift only one way or the other, we might be justified in expecting the local conditions to show few intermediate results. Closely allied to this is a second possibility, namely that in a particular *U*-shaped curve we really are not dealing with a simple curve, but with a compound one in which two opposite, extremely asymmetrical and apparently *J*-shaped curves have been combined.

As a third possibility, we have to note that since a *U*-shaped curve is merely one of the first or second curves upside down, it is possible that by taking the remainders or complements (especially if the data be in percentage form) of the dependent variable, we can right it again, taking the data out of the fourth class entirely and throwing it into the first or second (or even third) class, there to be treated as above outlined. And just as the symmetry of the minimum-maximum-minimum curve may be effected by the logarithmic projection, so the maximum-minimum-maximum or the *U*-shaped curve may be similarly converted from asymmetry to symmetry. Mortality rates often show an extremely asymmetrical *U*-shape, which can be changed to a valley of perfect symmetry by the log-scale, giving a beautifully rounded curve with clearly emphasised variations for different racial and occupational groups of the population.[6] And by subtracting the death rate in each age from the base of the rate, we obviously get what might be called a survival rate which, though less known, is a clear example of the second group of curves.

It is not to be understood from the foregoing discussion of statistical and graphical methods of producing symmetry in a curve, that all frequency series can be made symmetrical by proper treatment. Sometimes the very failure of the series to be symmetrical is of prime importance and while we might by round-about methods produce symmetry, yet the inappropriateness of these methods would be so great as to make symmetry meaningless. Nor is it true that all frequency curves will fall into one or another of the four classes mentioned. The point of what has been said in the foregoing discussion is that it is sometimes, indeed often, possible by very simple steps to

[6] The Census Bureau in its *Life Tables* has published elaborate charts of these curves, but unfortunately through the use of arithmetically-projected scales, it has been obliged to break up each curve into three or four parts, each part with a suitable but different, scale.

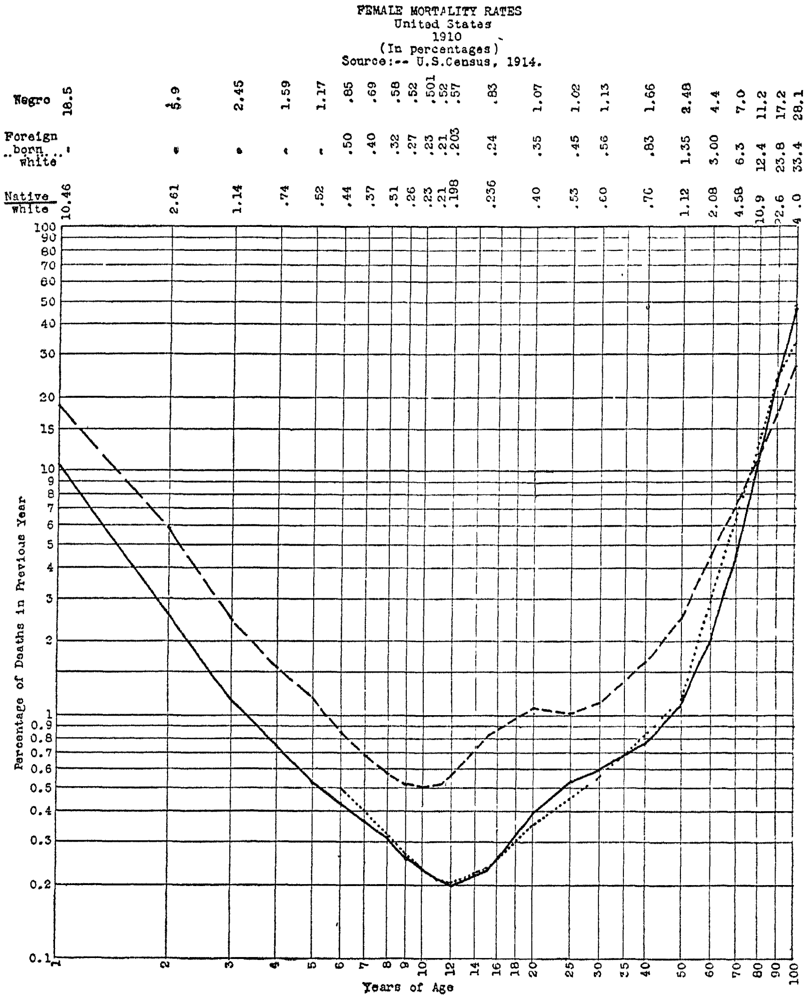


Fig. 363. An Extremely Asymmetrical U-shaped Distribution Brought to a Beautiful Symmetry by the Log-scales.

Note the emphasis given to significant irregularities, such as the increased mortality-rate at adolescence, among negroes, and its absence among the foreign-born.

attain the desired symmetry, and discover an underlying regularity in the behavior of the phenomena observed. To determine whether the symmetry is desirable and significant, and to interpret the meaning thereof when the symmetry has been secured, is indeed a task calling for experienced judgment. Furthermore, to detect the possibility of such symmetry

MORTALITY RATES
Deaths at Each Age, Shown as a Percentage of Those Living at the Beginning of the Previous Year
The United States
1901 and 1910
(Source:- U. S. Census Bureau)

1910	11.462	2.762	1.234	.783	.565	.466	.391	.330	.282	.247	.219	.236	.284	.352	.428	.542	.623	.772	.911	1.099	1.377	2.673	5.514	11.910	23.431	38.546
1901	12.448	3.352	1.567	1.016	.779	.605	.500	.413	.345	.297	.256	.277	.347	.439	.536	.690	.773	.877	.985	1.139	1.399	2.562	5.244	11.750	23.752	42.851

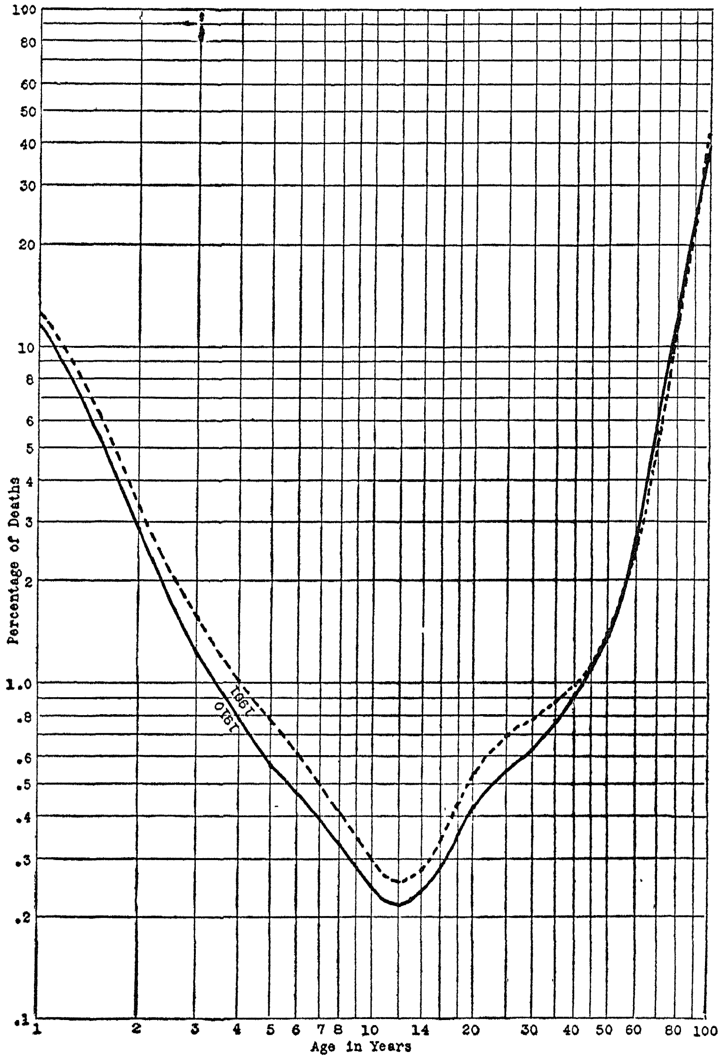


Fig. 364. Same as Previous—Historical Comparison.

requires imagination and familiarity with statistical data which the student will not quickly achieve.

Some idea of the sign posts which indicate the appropriateness of the logarithmic projection has already been given. When the class or group limits which break the phenomena into a series are at widely varying and rapidly increasing intervals, such as (0), 1, 2, 5, 10, 20, 50, 100, 200, . . . (infinity) or 1 month, 2 months, 3 months, 6 months, 1 year, 2 years, 5, 10, 15, 20, 25, 30, etc. years, and similar arrays, then the geometric

CONVENIENT GEOMETRIC INTERVALS

Number per "deck" (powers per ten)	2	3	(4)	5	6	7	8	10
Approximate Ratio Between Intervals	3 2	2 1	(1 8)	1 6	1.5	1 4	1 3	1 25
Intervals	1 3 10	1 2 5 10	1. 1.75 3. 5.5 10.	1. 1 5 2.5 4. 6. 10.	1. 1 4 2 3 4 5 6 5 10.	1. 1.4 2. 2.75 3 75 5 25 7.25 10.	1. 1 3 1 8 2 4 3.2 4 2 5.5 7.5 10.	1 1.25 1 6 2. 2.5 3 2 4. 5. 6.25 8. 10.

Fig. 365. A Table of the Convenient, Nearly-geometric Intervals by which the Range Between Successive Powers of Ten May be Divided.

nature of the progression is evident. But the law of organic growth applies in far more cases than those which wear this obvious marking. It may not be amiss to note some of Yule's illustrations, since we have followed his classification of curves. As he says, the symmetrical curve is rare in economic statistics and his one example of it, relating to the stature (height) of groups of adult men, covers so small a range in inches (from 58 to 76) that the logarithmic projection of the scale would not appreciably alter it. Hence we may conclude that while this particular series is practically symmetrical—on arithmetical projection—yet it may really call for a logarithmic projection because of the organic nature of human growth, and the arithmetical symmetry may be entirely accidental. This idea grows stronger as we observe that his illustrations of moderately

AMERICAN ACCIDENT TABLE
Duration of Temporary Total Disability
(95,388 cases per 100,000 accidents)
United States
1919
(Source:- Olive E. Outwater)

Weekly Rate

Period	Cases
0 - 1 week	44,673
1 - 2 weeks	17,712
2 - 3 "	10,925
3 - 4 "	6,269
4 - 5 "	4,345
5 - 6 "	2,674
6 - 7 "	1,923
7 - 8 "	1,298
8 - 9 "	966
9 - 10 "	745
10 - 11 "	549
11 - 12 "	447
12 - 13 "	358
13 - 14 "	296
14 - 15 "	264
15 - 16 "	218
16 - 17 "	182
17 - 18 "	154
18 - 19 "	131
19 - 20 "	110
20 - 21 "	94
21 - 22 "	73
22 - 23 "	68
23 - 24 "	56
24 - 25 "	45
Over 25 "	609

Daily Rate

Period	Cases
1 day	8,823
2 days	8,086
3 "	7,282
4 "	6,014
5 "	5,255
6 "	4,605
7 "	4,807
8 "	3,090
9 "	3,074
10 "	2,740
11 "	2,475
12 "	2,275
13 "	1,868
14 "	2,190

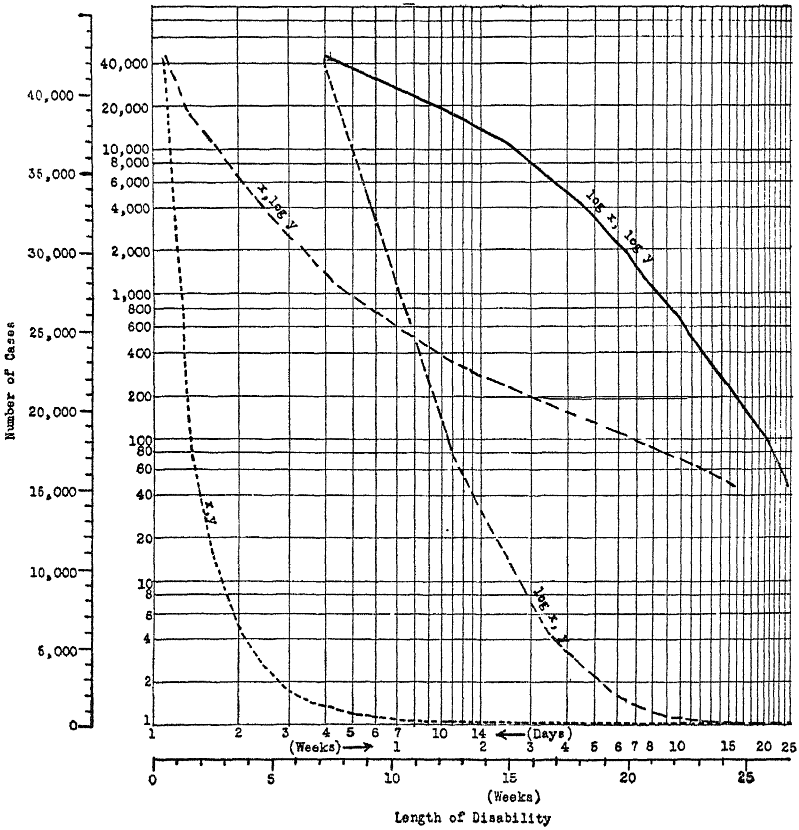


Fig. 366. The J-shaped Distribution.

(Shown by the dotted line and plotted upon the outer or arithmetic scales) becomes more rounded as logarithmic projection (shown by the full line and plotted by the two inner or logarithmic scales) is used.

asymmetrical curves include distributions of the stature of boys and young men, in which the ranges are considerably larger (great enough to show appreciable difference between the two projections) and in which the curve becomes symmetrical on a log-x scale. The weights of the adult men likewise showed moderate asymmetry which disappeared on the log-x scale, again because of a greater range (from 100 to 250 pounds). In all these cases, the class or group intervals are even and regular, and yet from the organic nature of the phenomenon—human growth in height or weight—we could suspect the desirability of the logarithmic projection which is so successful in fact. The true significance of the success of the log projection in producing symmetry, when it does so, is that the proper units of growth are magnitudes, such as those by which we measure star-brilliancy or musical pitch, rather than increments; nature uses geometrical, not arithmetical, units.

As has already been said and as may be deduced from the above, it is not always necessary that both axes should be upon the logarithmic projection. Frequently one has use for a chart, the scale of which is arithmetically projected along one axis and logarithmically along the other. Thus we may note at once four possible chart-fields: the plain rectilinear co-ordinates or x-arithmetic y-arithmetic; the logarithmic or x-log, y-log; and the two semi-logarithmic, x-arithmetic y-logarithmic, and x-logarithmic y-arithmetic. The discussion has so far turned upon the projection of the x-axis scale. With regard to the y-axis scale, the different projections obviously do not affect the symmetry of simple curves and we must base our selection of the proper projection either upon the nature of the variable to be shown and the apparent appropriateness of either method, or upon the emphasis or detail which we wish to give to certain parts of the curve, or upon mere convenience. The log projection of the y-scale always gives more pointed valleys and more rounded peaks than the arithmetical projection. Log-logs, or the logarithms of logarithms, have been mentioned in a previous chapter, and by their use still further changes can be effected in the contour of the curve. In short, the student will find ample means in these projections to study his data in various forms in the course of his analysis.

Chapter XXXVIII

LOGARITHMIC OGIVES

It is in regard to the ogive that the projection of the y-scale becomes important. It will be recalled from the chapter on ogives, that these charts show the cumulation of frequency series. It will also be recalled that the frequency series can be cumulated from either end, forming either a "more-than" or a "less-than" cumulation. Hence, even on an arithmetically projected field we can always have two ogives for the same original uncumulated frequency series. If this series be a symmetrical one, the two ogives will mirror each other on both axes; but if the original distribution is asymmetrical, they will not mirror both ways even when arithmetically projected. But by the use of logarithmic-scale projections, either on one or the other or both axes, we can sometimes produce mirroring again, a condition which often indicates that on such scales the curve would become symmetrical. For the treatment of the ogive, then, as for the simple curve, the proper projection of the x-axis is important. But in ogives we have, of course, only one maximum and one minimum, with the entire range of the series distributed between. Hence it is sometimes possible to secure in ogives what can never be secured in uncumulated frequency series—a straight line. The single exception to this is the J-shaped frequency curve with only one tail, and this curve will often, as has been said, be found to be an ogive in disguise, the cumulated nature of its data being not immediately apparent.

Now just as symmetry is more desirable in general than asymmetry, so a straight line is in general more desirable than a curve. For it still further simplifies the significance of the chart, and it still further displays regularity of behavior in the phenomena. Hence in work with ogives, which are merely irregular curves, typically of an S-shape, the search for a straight line is legitimate. And at this point, the value of the

log-projection of the y-axis scale comes in. For it may be that an ogive of very great curvature will straighten out into at

DURATION OF STRIKES
Number of Strikes ended in Specified Periods of Time or Less
United States
1916-1921
(*Note - Strikes omitted because lengths not reported:- 1916, 332; 1917, 616; 1918, 484; 1919, 301; 1920, 437, and 1921, 262.)
(Source - Monthly Labor Review)

Years	Total*																							
1921	1,147	67	99	142	185	217	249	290	317	376	404	442	511	553	569	599	630	664	714	751	828	885	940	1105
1920	1,262	88	152	205	255	291	333	396	441	501	552	609	691	714	751	803	846	864	907	955	1019	1066	1104	1221
1919	1,795	105	174	254	332	405	450	517	568	676	732	804	913	1007	1051	1116	1182	1240	1319	1393	1509	1579	1634	1776
1918	1,660	229	399	526	637	708	775	888	947	1041	1090	1166	1240	1307	1344	1376	1433	1461	1498	1530	1570	1586	1601	1629
1917	1,401	282	393	495	556	611	676	769	798	869	932	984	1058	1102	1124	1156	1184	1211	1248	1274	1311	1330	1341	1392
1916	2,063	179	362	508	632	762	871	962	1047	1201	1283	1370	1512	1594	1633	1693	1746	1771	1819	1841	1894	1933	1959	2046

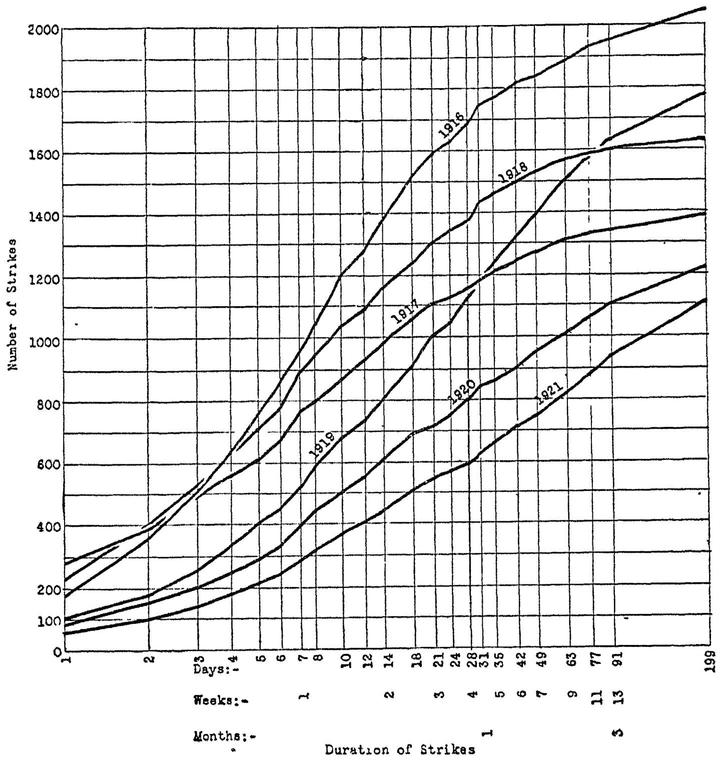


Fig. 367. **Ogives Plotted Upon Logarithmic Horizontal Scale.**
Compare Figs. 283 and 379.

least a close approximation of a straight line when the y-scale is made logarithmic. This possibility applies in general to all truly J-shaped curves. And when you have been unable to make a curve symmetrical, you still may succeed in making one of its ogives into a straight line, and so pull success out of defeat, and bring order where chaos was.

A spectacular example of this close approach to a straight line on the part of an ogive is the "more-than" cumulative of the distribution of personal incomes in a community. Government reports show tabulations of the number of persons who, according to their income tax statements, enjoyed incomes between specified limits. By cumulating this distribution so as to get the number of persons enjoying more than each specified amount of income, there is obtained the data for an ogive which will pass through such excessive ranges in both variables as a dozen persons enjoying more than ten millions of dollars annual income and two millions of persons enjoying more than a thousand dollars of income. The curve of this cumulated series will, on arithmetically-projected scales seem asymptote to both axes, hugging them so closely throughout its length that were the chart-field as large as the side of a house, the curve would never leave the axes by more than a few inches. But plot the same data on logarithmic paper and the ogive comes so close to a straight line that one is tempted to ascribe its variations to errors in the data. This particular example not only illustrates the tremendous compression of large numbers on a log-scale, but it also nicely exhibits the analytical power of the logarithmic method, for by its use the Italian economist, Vilfredo Pareto, was led to formulate a "law" of the distribution of wealth and income which was simply the mathematical expression of the slope of the straight line. And while Pareto's law has not withstood the waves of debate which it has occasioned, the straight line on which it was based remains the best means of analysis of comparative income statistics for different communities.

Another advantage in the logarithmic y-scale, and one which applies to all frequency curves, as well as to ogives, is that by its means, very dissimilar (as to amplitude or height on the y-scale) frequency curves can be compared. Thus when two series would on ordinary co-ordinates lie so far apart, the one high above the other, that they could not profitably be shown on the same chart without using different vertical scales

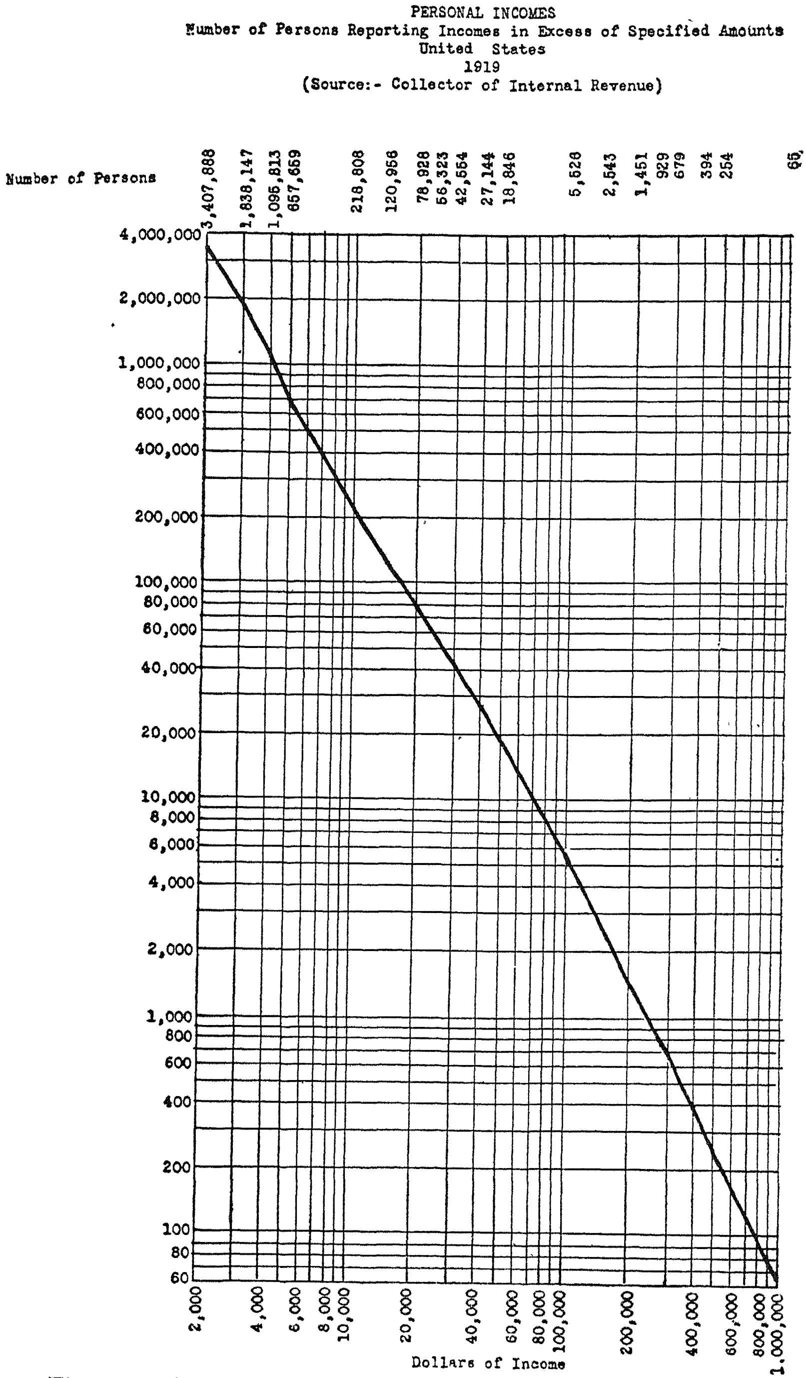


Fig. 368. An Ogive Plotted Upon Both Logarithmic Scales.

for them, it is often a very great saving in labor, as well as an assistance in analysis, to use the logarithmic vertical scales. The slopes of the various parts of the curves may be compared upon this projection and significance attached to parallelism, just as in historical rate of change curves. Commonly, the method is most useful when the two curves lie upon the same portion of the horizontal scale. Curves can also be made comparable by the use of percentages in the place of the numbers, each value being turned into a percentage of the total of the series. This requires more computing, but is for some purposes superior to the use of logarithms or logarithmic y-scale projection.

In the construction of the logarithmic frequency curves, the principles laid down under amount-of-change frequency curves apply as to the positioning of the field. When the ogive is used, there should be room above the field of the chart for the original data to enable reading from the independent variable, and it may often be well to leave room to the right of the chart for derived secondary data in the form of readings from the dependent variable. As for other frequency curves, the selection of the independent variable is sometimes difficult and often merely a matter of whim, choice, or convenience. In such cases, particularly, both the original and the secondary derived data are useful, for you cannot be sure in advance which data you will ultimately adopt. There is, however, in the logarithmic frequency curve, little use for the staircase form of plotting, as the areas between ordinates have no significance, and we can limit the discussion of the logarithmic curve to frequency polygons and smoothed forms of curves. As to the logarithmic projection of scales, of course, the principles laid down in the previous chapters apply. And other things being equal, the logarithmic projection of the x- and y-axes of each chart should be upon a common scale, that is, with decks of equal size, whatever the calibration may be.

The possibilities of application of the logarithmic frequency curve and the logarithmic ogive are very great and the student will soon discover that they exceed in usefulness for research purposes the ordinary amount-of-change frequency curves as much as the historical rate-of-change curves exceed the historical amount-of-change curves.

We have mentioned population distributions as obviously calling for logarithmic projection. This is important in sales

analysis and merchandising research. Rent statistics, possibly an even better index of local buying power than income statistics, have been found, where they have been compiled, to behave as do the incomes. In building statistics, it has been found possible to set up normals of new building for each community on the basis of size of population, and the comparison of the actual building with this normal affords a useful index of local business conditions, the entire analysis being carried out on logarithmic paper. The possible useful applications of this form of chart are innumerable. In the engineering world it is in common use, being much better known than the semi-logarithmic, historical rate-of-change curves. Though less popular than the latter and perhaps less often required, the logarithmic frequency curve should play an important role in the business or economic research laboratory.

PART IV. SPECIAL ANALYSES

Chapter XXXIX

THE NORMAL CURVE OF ERROR

The statistical authorities usually make a great to-do over frequency curves, for much statistical work of a precise nature involves the close study and measurement of frequency distributions. The fundamental conception in such work is one ideal or theoretical form of distribution which is known as the "normal curve." The analysis is more often than not directed at the question of whether given distributions conform to this normal, and if not, how closely they approximate it. In graphics the question is whether a given frequency curve can be found to be identical or nearly identical with a corresponding normal curve. For reasons which will presently appear, the discovery of this identity is always attended by rejoicing and relief, similar to the discovery that a historical curve conforms to the law of organic growth. For in each case a condition of meaningless irregularity has been replaced by the establishment of a definite and significant law.

If you select at random a hundred men, provide them with yard-sticks, and let them measure the length of the city block in front of your house, you will be disappointed if you expect a unanimous report from them. Their measurements will vary through a considerable range, and "bunch up" most thickly about midway between the extremes. Pick your men for experience and ability and you may expect the variation to extend through a shorter range, that is, the highest and lowest estimates will be nearer together, but the variation will still be present, with its bunching up at the mid-point. Provide the men with accurate engineers' chains instead of wooden yard-sticks and you will still further reduce the range of variation, perhaps down to inches instead of feet or yards, but it will still be present. The results are not changed if, instead of a hundred men, you use but one man, letting him measure the distance a hundred times over (provided he does not

voluntarily repeat his estimates). As a matter of fact, all these measurements are approximations; the actual distance itself has not changed or varied in the least. We are simply confronted with the human equation and its inevitable errors. And the interesting thing is that these mistakes or errors have been found to group themselves in a certain characteristic formation or distribution. In any sufficiently large body of observations the scatteration or dispersion of items due to mistakes or errors of observation, falls ever into or approximates the same characteristic form. Plotted upon a chart, this form is the normal curve, and the normal curve is therefore often called the curve of error or the curve of errors.

Under the name of "the probable error," artillery officers study the scattering of gun-fire, for no two successive shots from the best cannon in the world will hit at precisely the same spot. If you step outdoors and pick (without choosing) a hundred leaves from a tree you will find that while they may be of approximately uniform size, yet there will be minute variations of length or breadth in these leaves. Or glance at the stock exchange quotations and arrange the first hundred into a frequency series. In every case the form of normal curve will again be approached by the curves of your observations. Hence the normal curve is often called the "probable curve" or the curve of normal probabilities.

In algebra every school boy knows the expansion of the binomial $(a+b)$. The square of the binomial is $(a^2+2ab+b^2)$. Its cube is $(a^3+3a^2b+3ab^2+b^3)$, its fourth power is $(a^4+4a^3b+6a^2b^2+4ab^3+b^4)$; its fifth is $(a^5+5a^4b+10a^3b^2+10a^2b^3+5ab^4+b^5)$; and so on. Note that these coefficients increase always as they approach the center from either end. If we were to plot them as a frequency curve we should again find a suggestion of the normal curve. Carry the expansion out into higher powers and the approximation becomes closer. If the expansion could be made indefinitely great, the curve of the coefficients would precisely conform to the normal curve and for this reason the statistical study of the normal curve is closely tied up with the binomial theorem. If in no other way, the normal curve could be mathematically computed by its means.

Enough has been said to show the importance of the normal frequency distribution. Let us therefore attempt to describe its appearance. This is not an easy task, in as much as any

THE NORMAL CURVE

Ordinates of the Uncumulated Series

(Formula: $y = e^{-\frac{1}{2}x^2}$)

(NOTE: Values (ordinates) are given for parts of the range (abscissae) on both sides of the Median (origin), the units of measurement for abscissae being the Standard Deviation, and for ordinates, the Median.)

Negative Side of Median (origin)		Positive Side of Median (origin)
Abscissa (x)	Ordinate (y)	Abscissa (x)
0	1.000,0	0
—0.2	.980,2	0.2
—0.4	.923,1	0.4
—0.6	.835,3	0.6
—0.8	.726,2	0.8
—1.0	.606,5	1.0
—1.2	.486,8	1.2
—1.4	.375,3	1.4
—1.6	.278,0	1.6
—1.8	.197,90	1.8
—2.0	.135,34	2.0
—2.2	.088,92	2.2
—2.4	.056,14	2.4
—2.6	.034,05	2.6
—2.8	.019,84	2.8
—3.0	.011,109	3.0
—3.2	.005,976	3.2
—3.4	.003,089	3.4
—3.6	.001,533,8	3.6
—3.8	.000,731,8	3.8
—4.0	.000,335,5	4.0
—4.2	.000,147,75	4.2
—4.4	.000,062,52	4.4
—4.6	.000,025,42	4.6
—4 8	.000,009,930	4.8
—5.0	.000,003,727	5.0

Fig. 369.

change of scales upon the chart, along either axis, results in a change of its shape. Enlarge the horizontal scale and the figure of the curve is spread out wider, reduce that scale and it becomes narrower. Increase the vertical scale and every part of the curve is raised, making its peak much taller; diminish the scale and it is flattened out. Hence the normal curve may have an infinite number of conformations. Always, however, it has a hump at its horizontal mid-point where the fre-

quencies are thickest. Hence it may be described as bell-shaped, showing a peak at the mid-points and a die-away curve or tail at each side of the peak, one at least of the tails being asymptote to (i.e. approaching but never reaching) the horizontal axis.

Now there are any number of possible curves which fit the above description but are not normal curves, hence the description is not a definition. Curves may be rounder or flatter or more pointed at the top or may slope away at different angles from the corresponding normal curve or any similar normal curves. Hence it is important to be able to distinguish between normal curves and curves which are not normal. It is not enough to find in analyzing a given distribution, that its curve has a central peak and is symmetrical. The student should be familiar with the various forms of the normal curve, but it is not always possible, even for an expert, to be quite sure whether a given curve is normal or not, if he can rely only upon inspection. We need to compare the given curve with a precise drawing of the corresponding normal curve. And remembering the infinite number of normal curves, it becomes difficult to select the proper one for the given distribution. Obviously the compared normal curve, that is the ideal distribution, for a given series, is one which will have the same total area under the curve, that is, the same number of observations or items in the series. But there is an infinite number of normal curves with the same inscribed area, differing from each other according as their peaks are tall and narrow or short and broad. So the proper normal for any series must be one which will intersect the given curve at certain points. Usually the standard deviation is used as the abscissae of these intersecting points. So we are led into a tedious mathematical process, only a part of which is the computing of what statisticians call the standard deviation and all of which is more statistical than graphic. The comparison of a given curve and its normal is not easy by this method.[1]

In the next chapter will be described an easy trick by which you can make such a comparison graphically, and learn whether the given distribution or series is normal or not, and if not, how closely it approximates the normal distribution.

[1] The two ways of fitting the proper normal to a given curve, other than the graphic one in the following chapter, can be found in Yule, *Theory of Statistics*, pp. 307–9.

Chapter XL

PROBABILITY CURVES

Those readers who have understood the form of graphic legerdemain by which the adherence of a historical series to the law of organic growth is flashed upon the rate-of-change chart paper, will be prepared for a similar trick by which the adherence of a frequency series to the normal distribution is graphically shown. They will anticipate that the graphic form eliminates practically all computing and calculating from the treatment of the data, this computing having been absorbed once for all into the projection of the scale of the chart. The trick is very simple. It consists of plotting the ogive of the distribution upon paper which has been specially ruled off in such a way that the ogive of any normal curve will become a straight line upon it.

The normal ogive, as you know, is S-shaped. If you know the ordinates of the normal ogive you know precisely how much to distort the scale for the dependent variable so as to produce a straight-line projection of the normal ogive. You need only select points equidistant along the vertical scale, read the abscissae of corresponding points on the ogive, lay off these horizontal values vertically (on the vertical scale), and shift the scale figures from the old to the new vertical scale. By doing this you have made the ordinates and the abscissae of each point on the normal curve alike and so of course the normal curve becomes a straight line. But any other S-shaped curve, any other ogive, or any curve at all, which is not a normal one, will fail to straighten out perfectly. And so at a glance you can see, by this chart, not only whether a given distribution is normal, but if it is not normal, how closely it approximates or deviates from a normal distribution.

To plot a given ogive upon the probabilities projection of the dependent-variable scale just described, it is best to turn all frequencies, i.e. dependent variables in the data, into per-

Ordinates of the Cumulated Series

(NOTE: Values (ordinates) are given for parts of the range (abscissae) measured in both directions from the Median as origin, the units of measurement for abscissae being the Standard Deviation, and for ordinates, the total of the series. The table can also be used to show fractions of the area under the normal curve (uncumulated) lying on each side of verticals (ordinates) from specified abscissae.)

Negative Side of Median (origin)	Cumulation		Positive Side of Median (origin)
	"Less than"	"More than"	
	"More than"	"Less than"	
Abscissa x	Ordinate y		Abscissa x
0	.500,0	.500,0	0
—.05	.480,1	.519,9	.05
—.10	.460,2	.539,8	.10
—.15	.440,4	.559,6	.15
—.20	.420,7	.579,3	.20
—.25	.401,3	.598,7	.25
—.30	.382,1	.617,9	.30
—.35	.363,2	.636,8	.35
—.40	.344,6	.655,4	.40
—.45	.326,4	.673,6	.45
—.50	.308,5	.691,5	.50
—.55	.291,2	.708,8	.55
—.60	.274,3	.725,7	.60
—.65	.257,8	.742,2	.65
—.70	.242,0	.758,0	.70
—.75	.226,6	.773,4	.75
—.80	.211,9	.788,1	.80
—.85	.197,7	.802,3	.85
—.90	.184,1	.815,9	.90
—.95	.171,1	.828,9	.95
—1.0	.158,66	.841,34	1 0
—1.1	.135,67	.864,33	1.1
—1.2	.115,07	.884,93	1.2
—1.3	.096,80	.903,20	1.3
—1.4	.080,76	.919,24	1.4
—1.5	.066,81	.933,19	1.5
—1.6	.054,80	.945,20	1.6
—1.7	.044,57	.955,43	1.7
—1.8	.035,93	.964,07	1.8
—1.9	.028,72	.971,28	1.9
—2.0	.022,75	.977,25	2.0
—2.2	.013,90	.986,10	2.2
—2.4	.008,20	.991,80	2.4
—2.6	.004,66	.995,34	2.6
—2.8	.002,56	.997,44	2.8
—3.0	.001,35	.998,65	3.0
—3.2	.000,69	.999,31	3.2
—3.4	.000,34	.999,66	3.4
—3.6	.000,159	.999,841	3.6
—3.8	.000,072	.999,928	3.8
—4.0	.000,032	.999,968	4.0
—4.5	.000.003	.999,997	4.5

Fig. 370.

centage figures. For the peculiar spacings of the probabilities projection are arbitrary and cannot be freely changed. The total of a series is the limit of its cumulation, and the cumu-

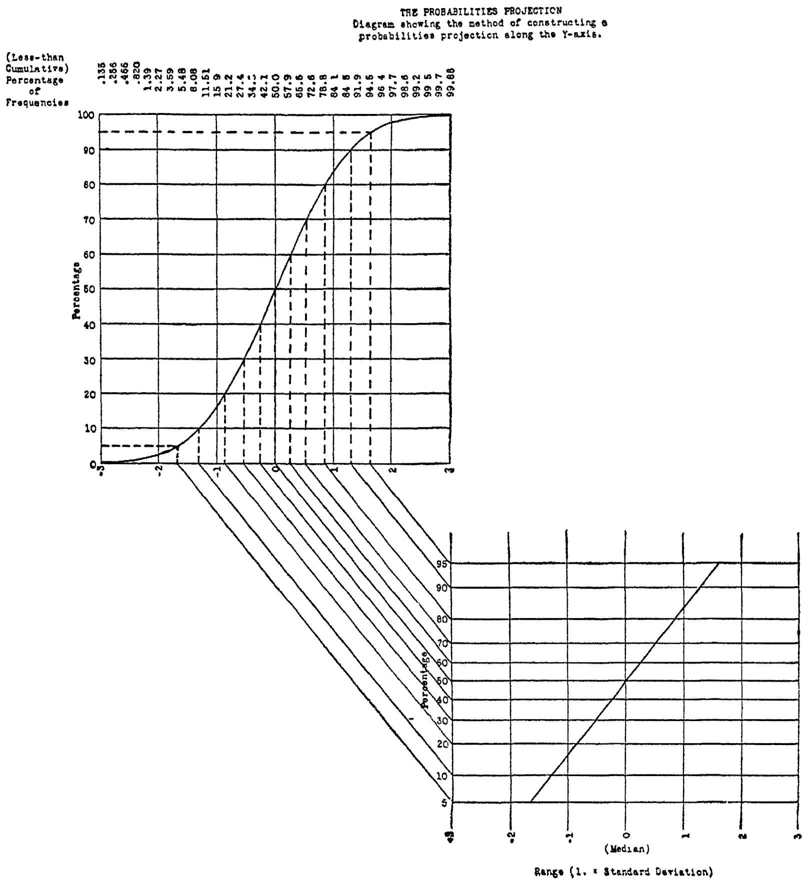


Fig. 371.

Showing how the dependent y-scale of percentage-frequencies is readjusted from the arithmetical projection which gives the curved ogive to the probabilities projection which gives the straight-line ogive. The frequencies must invariably be converted into percentages for this chart.

lated series has definite limits (for its dependent variable), these limits being "no items" and "all items," that is, 0 and 100% of the total of the series. Hence the only common measure for all frequency series is a percentage one and the normal curve and ogive are both given in all tables in per-

centages. The distortion of the dependent variable scale is made for values of these percentages, and you must not seek to shift the scale figures of the probabilities projection scale as you could an arithmetical or logarithmical projection scale. Your only way to alter the probabilities scale is to turn the percentages into values of your particular series, in which case the scale is restricted to this series, a fruitless if not dangerous step. It is sufficient to turn the cumulated series into percentages before plotting upon the probabilities projection.

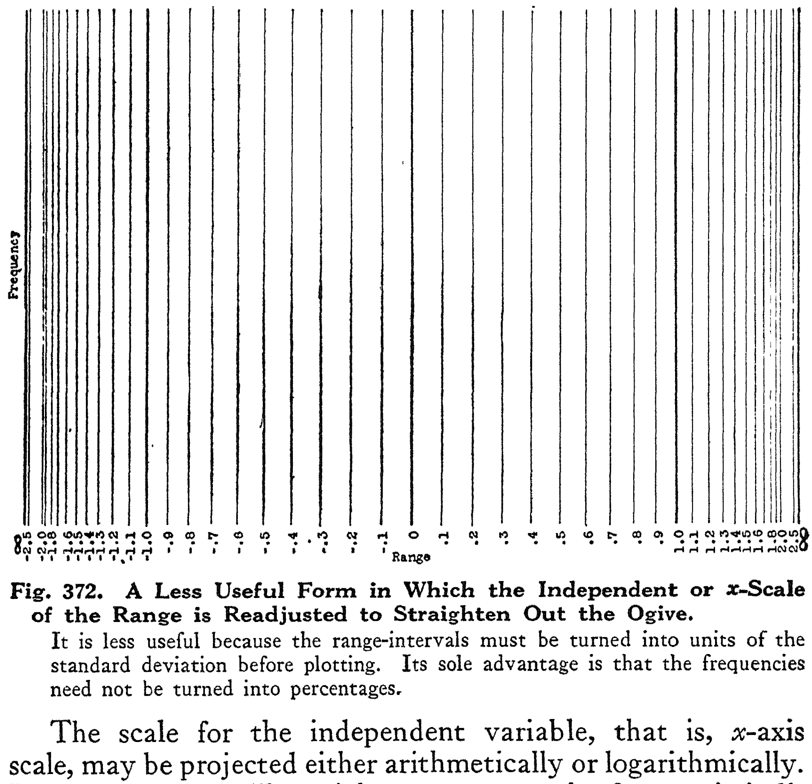


Fig. 372. A Less Useful Form in Which the Independent or x-Scale of the Range is Readjusted to Straighten Out the Ogive.

It is less useful because the range-intervals must be turned into units of the standard deviation before plotting. Its sole advantage is that the frequencies need not be turned into percentages.

The scale for the independent variable, that is, x-axis scale, may be projected either arithmetically or logarithmically. If a given series will straighten out upon the former, it indicates that its curve would be symmetrical upon an arithmetic projection; if the ogive straightens out upon the logarithmic paper the distribution is one of the asymmetrical types which are made symmetrical by a logarithmic projection. Hence the probabilities paper not only gives the approximation to the normal, but it also determines whether that normal be symmetrical upon a geometrical or arithmetical basis. Chart

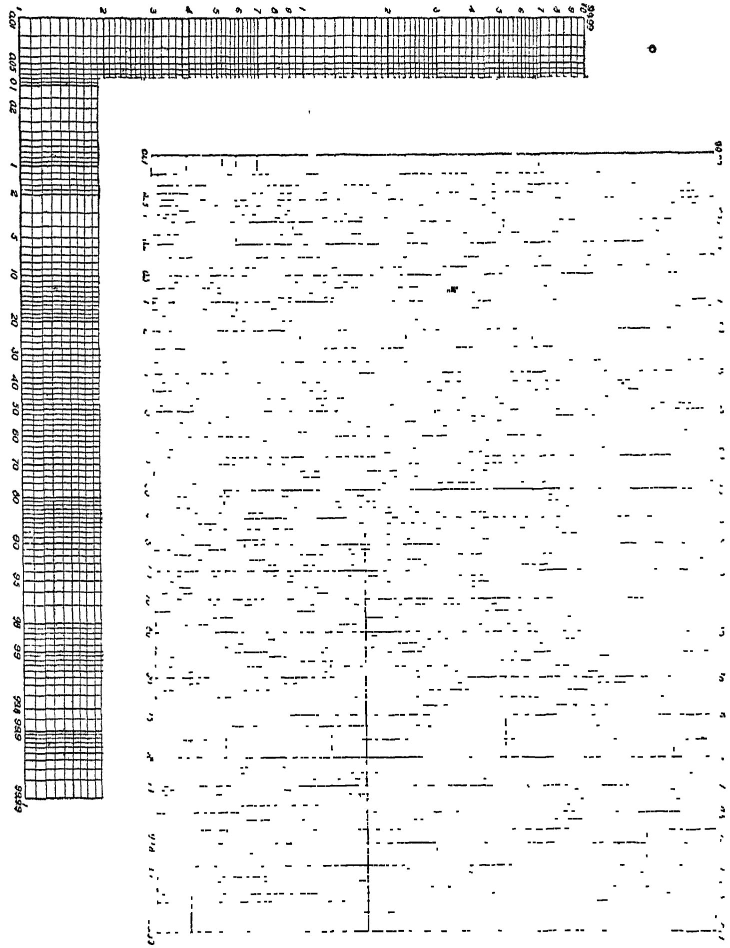


Fig. 373. Commercial Probability Forms.

These two sheets, an arithmetical-probability form and a logarithmic-probability form, are marketed under these names by the Codex Book Company.—*Permission of the Codex Book Co.*

forms with the probabilities projection of the dependent variable scale, are published with both projections of the independent variable scale.[1]

[1] The true nature of the "dependent" and "independent" variables in ogives has already been discussed. (See Chapter XXIX.) The same considerations hold of the ogive on probabilities paper, but by accident some of the publishers of probabilities paper have reversed the arrangement in their printed scales, thus inadvertently hastening the time when, in the author's opinion, statistical practice will correct itself and plot x-frequencies, y-range.

The probabilities scale, calibrated in percentages, will never reach either zero or one hundred per cent in either direction,

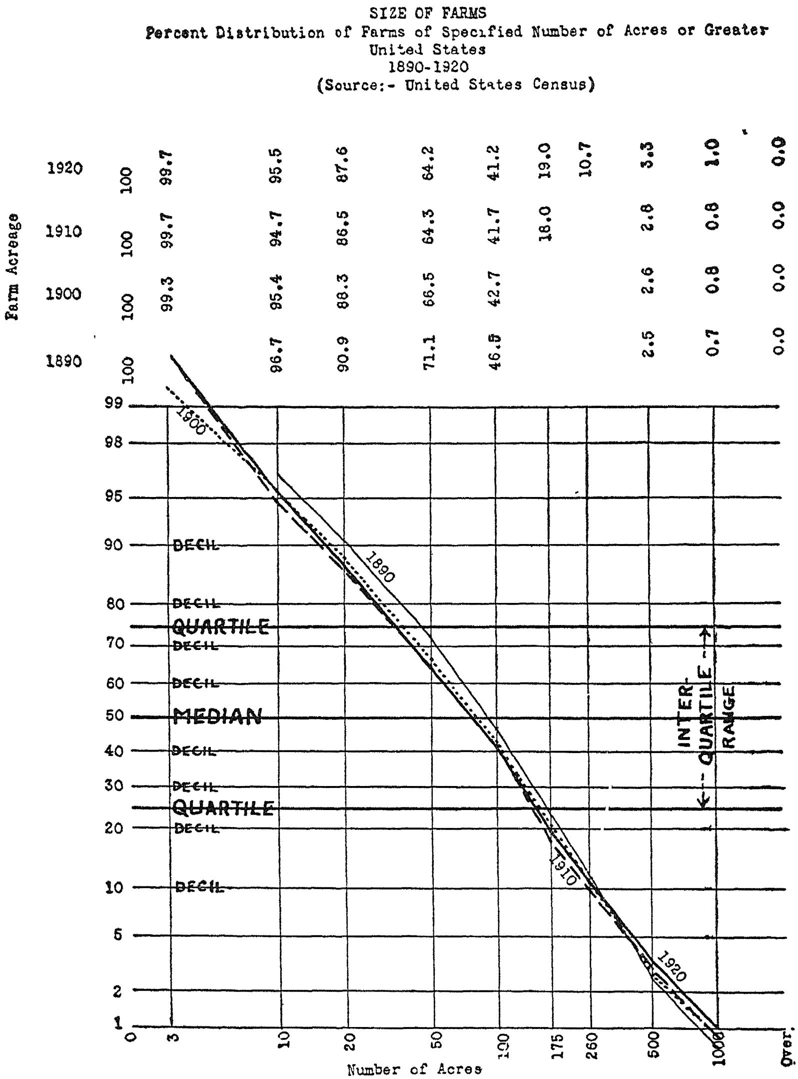


Fig. 374. The Logarithmic Probabilities Projection in Use.
Note the close approach to straight lines. Also the ease with which, median, quartiles, etc., and interquartile range are found.

for the true normal curve is asymptote to the x-axis and its 100% parallel. We can therefore never plot the two limits,

0 and 100%, on this paper. We can only come as close as we wish toward these limits by extending the scale on into the

COLLEGE SALARIES
Salaries in American Colleges and Universities
Including Public and Private Institutions
United States
1920
(Source - United States Bureau of Education)

Salary	(15)	(25)	(40)	(60)	(100)	(150)	(250)	(400)	(600)	(1000)	(1500)	(2500)	(4000)	(6000)	(10000)	(15000)
Presidents									100,000	99.725	98.62	81.5	47.6	20.6	5.6	0
Deans and Directors							100,000	99.94	99.65	97.3	88.3	55.8	19.1	3.87	.17	0
Full Professors				100,000	99.992	99.968	99.928	99.62	99.17	95.0	83.5	46.6	13.8	.96	.018	0
Associate Professors							100,000	99.915	99.87	99.40	96.0	47.4	.72	0		
Assistant Professors				100,000	99.93	99.88	99.81	99.61	98.35	94.2	79.3	14.9	.125	0		
Instructors			100,000	99.75	99.41	99.11	98.50	97.15	94.3	82.6	49.5	1.13	.04	0		
Assistants	99.93	99.83	99.93	99.16	93.38	89.52	77.97	68.25	51.95	24.60	6.60	.10	0			

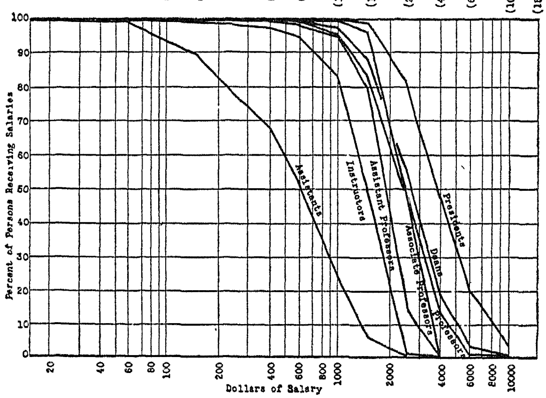


Fig. 375. An Arithmetical Projection of the Dependent (or Frequency) Scale.

small fractions near each limit. After we pass the points of .01% and 99.99% the scale resembles a logarithmic projection so closely that we can add to it by logarithmic scales. But the tails of a curve are least significant, and it is rarely worth while to make this extension. Unless we deal with large series containing over ten thousand items and so grouped that the terminal groups have only one item each, we shall not need to plot points less than .01% or greater than 99.99%. Usually

it is safe to chop off all of the scale beyond 1% and 99%, thus reducing the chart, either because of the absence of data out-

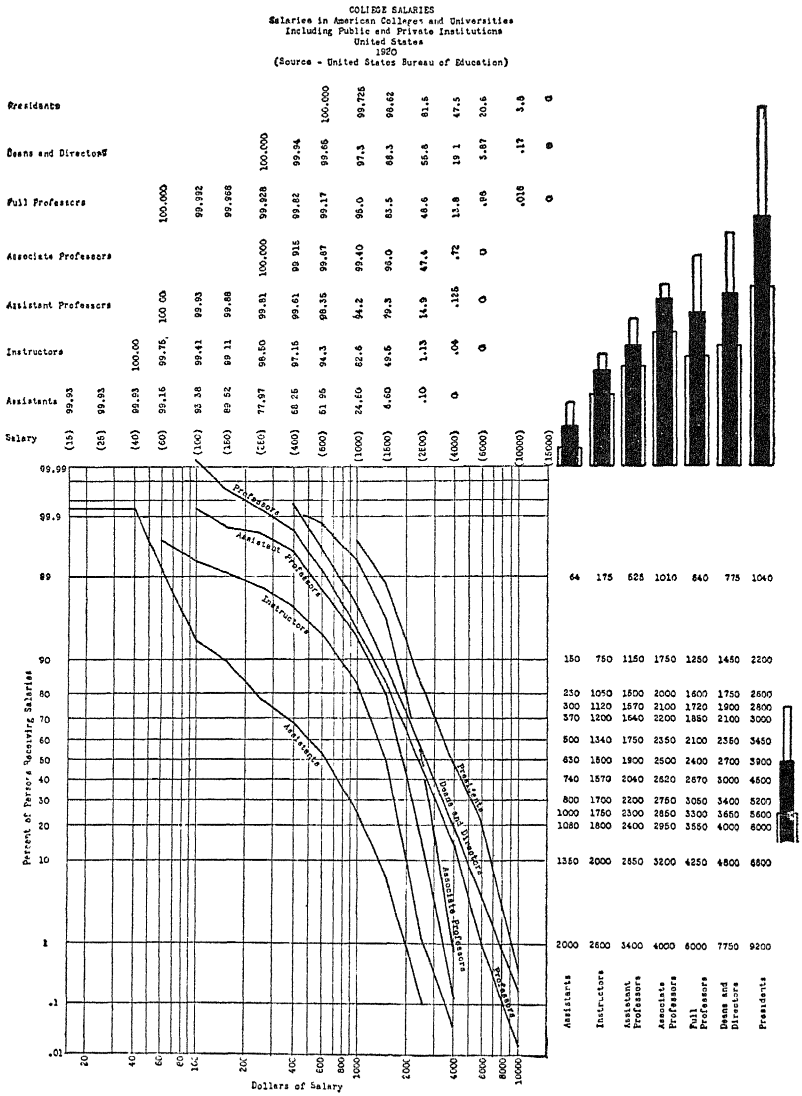


Fig. 376. A Probabilities Projection of the Previous Chart.

Note the secondary data and bar-chart obtained from interpolation of the curve at the various decils, showing the median and other salaries for each group of educators. The swing of the upper tails of the ogives away from the straight line of the normal distribution might be interpreted as due to a few institutions which give titles out of proportion to the salaries attached thereto.

side this range (other than the limits) or because of the lack of their significance.

The significance of the ogive is always elusive to the layman and the probabilities projection of the ogive, while it clears up, for the technician, the question of normality, is even more baffling in other respects. A little study will show us, however, that both the direction of the curve and its position are significant. When comparing two ogives, if we find the ogive of one series further out along the horizontal scale than the

OUTPUT OF FACTORIES

The Value of Products of Manufacturing Establishments having less than specified value of products per establishment
United States
1904-1914
In percentages of the total
Source:--U S Census

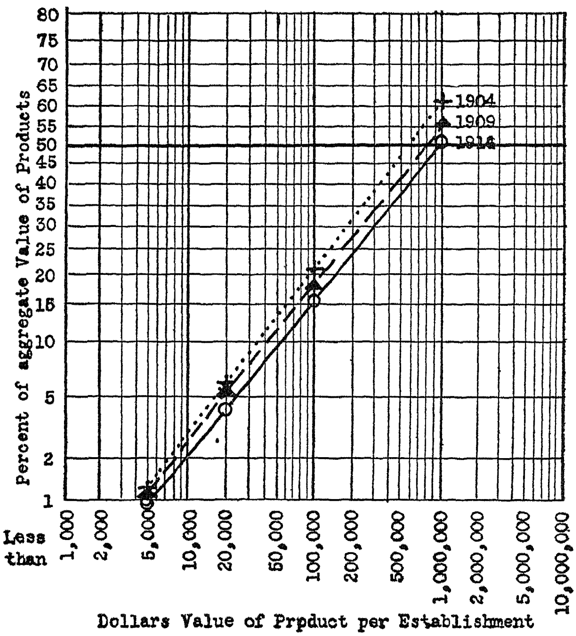


Fig. 377. Symmetrical so Far as Data Obtains.

other, it means that the items in that series are greater than corresponding items (in similar parts of the distribution) in the other series. Thus an ogive for the heights of children

will be to the left of an ogive for the heights of adults. If the two simple frequency curves had been plotted they might overlap but the main bodies of each would be at different positions along the scale. The significance of the positions of ogives will be clear if the corresponding simple curves be

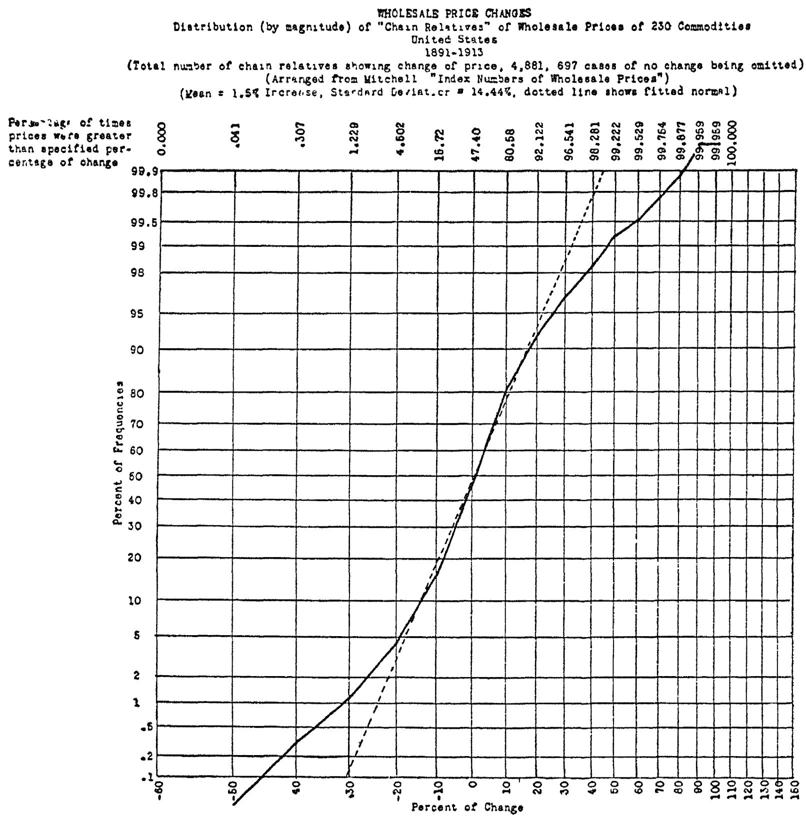


Fig. 378. Symmetrical but Distinctly not Normal.

imagined underneath them or if horizonal bars be imagined as lying between each ogive and the y-axis.

As in other ogive-charts, secondary data may be derived from the curves, being the reading upon the x-axis scale of intersepts of the curve and the horizontal rulings. In other words, while the curve is drawn by plotting the ordinates at certain abscissae we may interpolate from it the abscissae of certain ordinates. The readings are usually taken at the

median and quartiles, and decils, and occasionally at the extreme percentiles.[2] In this derived or secondary data we see a reversal of the dependence of the variables, the dependent one

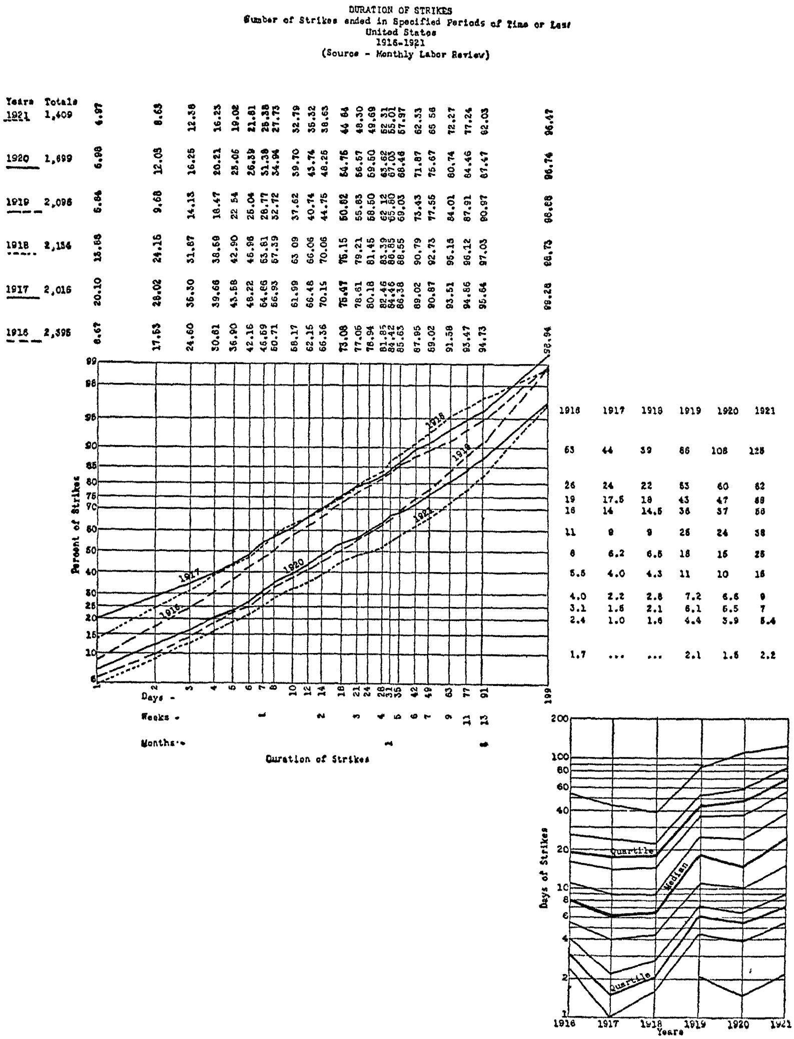


Fig. 379. The Comparison of Ogives for Different Dates.
This gives a historical curve instead of a bar-chart for the secondary derived data. The ogives are useful for interpolation but the derived curves afford the best view of changes.

---

[2] That the interpolation for median, quartiles, etc., does not give cases, but values, has already been pointed out. (See Chapter XXIX.)

becoming independent. This secondary data affords most of the statistical measures of both dispersion and skew.

The significance of the direction of the ogive upon the probabilities paper (when it approximates a straight line) is more subtle, but far-reaching. For as we have seen, the range of observations may vary in different series. In the example already given of measurements of a given distance, we saw that as precision of measurement is increased, the distribution becomes more concentrated, the scatteration less, and the peak of the simple curve taller. It is precisely this condition which will make the ogive curve more nearly perpendicular to the x-axis. As the dispersion increases and the precision diminishes, the ogive curve swings about toward a horizontal direction. These considerations hold as well for the ordinary ogive, but are more clearly seen and measurable in the ogive projected upon probabilities paper.[3]

In the analysis of frequency data by means of the ogive curve upon the probabilities projection, a difficulty is often met in that the data is incomplete, no figures being obtainable for a portion of the range. Thus the statistics of income do not include the personal incomes below two thousand dollars for heads of families and one thousand dollars for individuals, and for this reason omit perhaps ninety per cent of the population. Astronomers have estimated the number of stars of each magnitude down to the twentieth, but do not carry their estimates much further. In such cases as these we do not know the entire "population," "universe," or total body for which the distribution applies, and hence cannot turn frequencies into percentages.

The problem is largely statistical but affects the subject of charts in that an ogive upon probabilities paper, of these incomplete distributions, obtained by turning the frequencies into percentages of the known sub-total, will almost certainly fail to form a straight line although the entire distribution may be perfectly normal. It is not legitimate to plot such parts of distributions upon the probabilities paper unless we can turn the frequencies into percentages of the true total. At this point, therefore, we will mention a statistical trick by which you can sometimes dodge the difficulty. For in all frequency data there are two possible series, both covering the

[3] On the probabilities projection, the mode can not be graphically determined by the slope of the curve, as it could on arithmetical projections of the ogive.

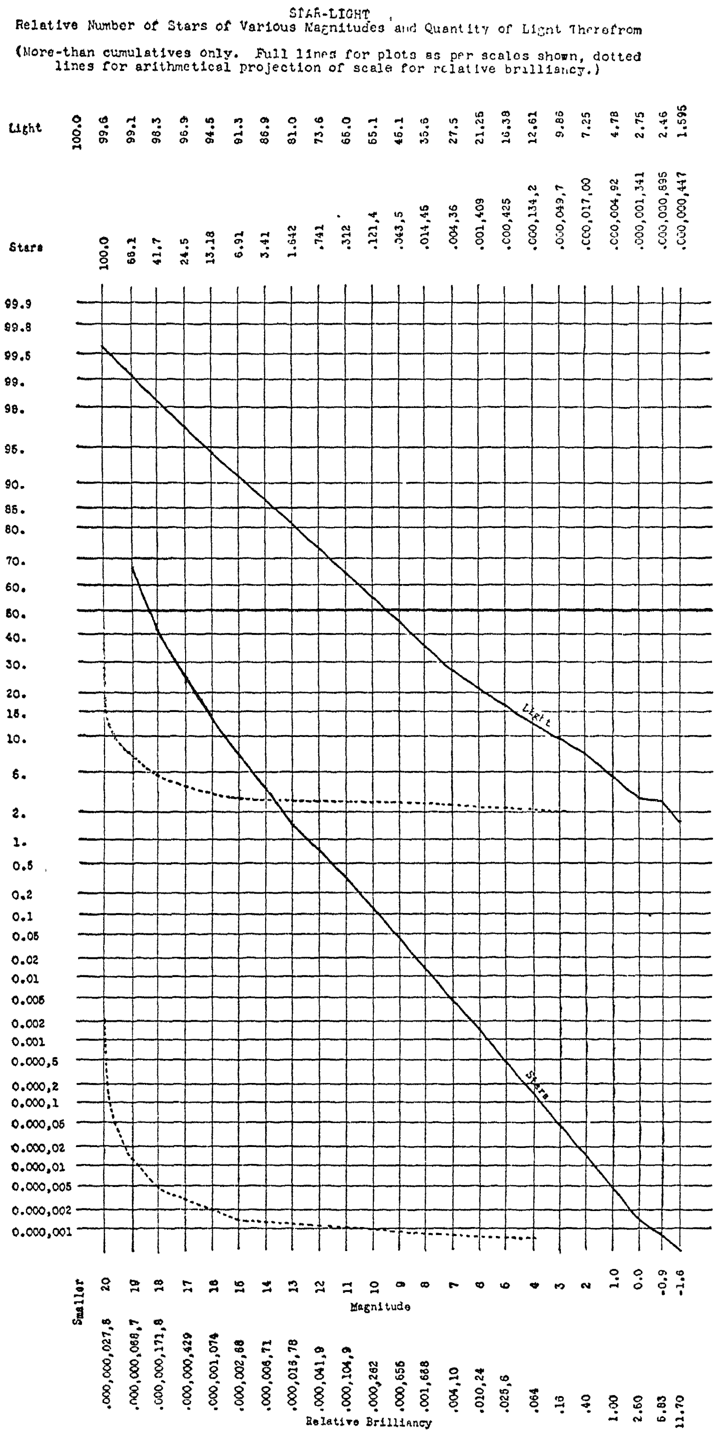


Fig. 380. The Ogive of the Units of Measurement of the Items (Star Brilliancy) in an Incomplete Series is Straighter and more Reliable than the Ogive of the Items (Number of Stars).

See footnote on opposite page.

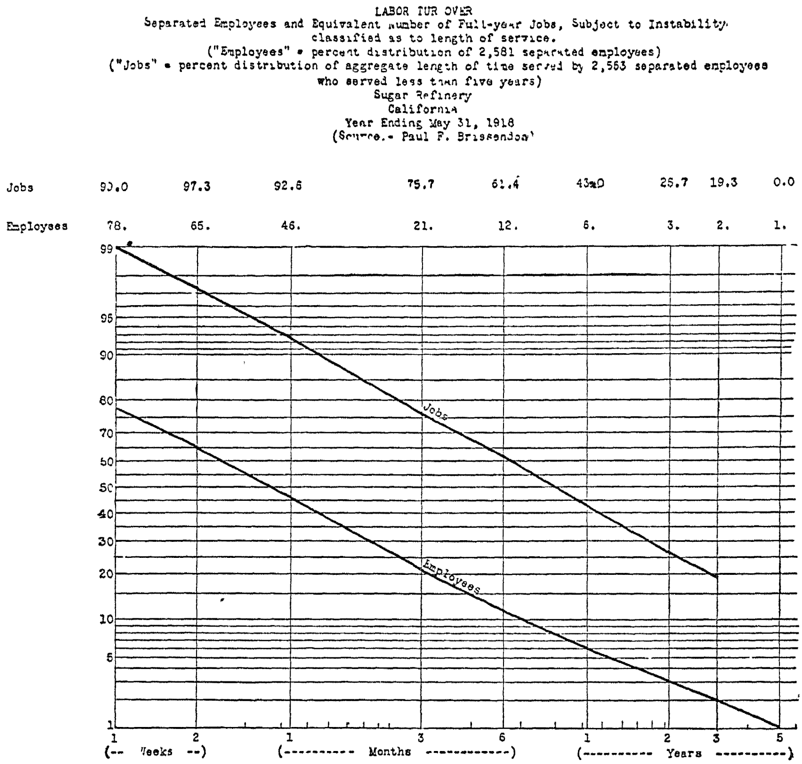


Fig. 381. Another Example of the Two Interconvertible Frequencies for the Same Data.

same observations. These two series have been described in an earlier chapter. They are, briefly, the count of items, and the count of units of measurement of the items, group by group, through the distribution. The point is that these two series for the same phenomenon are often available and can usually be estimated if not available. And if a considerable portion of the data of one series be missing, it is generally true that the missing portion covered items of small or negligible numbers of units individually. If then we convert the

---

NOTE TO FIG. 380

The dotted lines are plotted upon an arithmetical scale (not shown on the chart) of relative brilliancy. The full lines are plotted by the two horizontal scales (appearing on the chart), namely a logarithmic projection of relative brilliancy and an equivalent arithmetical projection of magnitudes. It would also have been possible to project the magnitudes logarithmically (thus obtaining a log-log projection of brilliancies). The interesting point is that star magnitudes though projected arithmetically are in themselves geometric units and form a logarithmic projection of brilliancies.

item data into unit data, we generally find that the importance of the missing portion of the data has greatly diminished. Thus while income statistics omit perhaps ninety per cent of the families and individuals in the country, they omit only about ten per cent of the total income. In the case of stars, the astronomers have computed the light of all stars, so that the unit data is complete, while the item-data is incomplete. When the incomplete data can be reduced to a relatively small amount in unit-data, it may more safely be estimated, and so completed. Thus by converting the data into another form, we may find it possible to project the ogive curve upon probabilities paper with satisfactory results.[4]

OUTPUT OF FACTORIES

The Value of Products of groups of Employees
employed in Manufacturing Establishments
(all groups composed of employees in establishments
having lowest value of products)
United States
1904 - 1914
(In percentages of total)
Source: U S Census

Class of Establishment by Value of Products	1904		1909		1914	
	Employees Number	Product Value	Employees Number	Product Value	Employees Number	Product Value
Less than $5,000	1.9	1.2	2.2	1.1	1.8	1.0
Less than $20,000	9.6	6.3	9.3	5.5	7.9	4.7
Less than $100,000	28.4	20.7	25.8	17.8	22.1	15.2
Less than $1,000,000	74.4	62.0	69.6	56.2	64.8	51.3
Any value whatever	100.0	100.0	100.0	100.0	100.0	100.0

Fig. 382. Alternative Data Yielding the Lorenz Curves.

[4] "If the observations are not complete (i.e. cover only a small part of the unknown total range of the variable, though it is highly desirable that what observations we have do not constitute a mere extreme tail), it is possible to fit a normal curve to the data by means of fitting a second degree parabola, by the method of least squares ($y=A+Bx+Cx^2$) to the logarithms of the number of observations in each interval. The theory is based upon the equation of the normal curve,

Taking logs of both sides, $y=K_1 e^{\frac{-x^2}{K_2}}$

$$\log_e y = \log_e K_1 - \frac{1}{K_2} x^2$$

$$= K_3 - K_4 x^2$$

we get a second-degree parabola."—Dr. Frederick R. Macauly.

We have spoken of the alternative series into which any distribution may be converted. The combination of these two yields, upon arithmetical paper, the Lorenz curve. And it is for the Lorenz curve, of all types, that we can profitably use double-probabilities paper, that is, paper projected upon the probabilities scale along both axes. For, as will be remembered, the tails of a Lorenz curve are nearly asymptote to the axes when the dispersion is great, and the values near either extreme become difficult to interpolate or read from the chart. Moreover, when the groups or classes, into which the distribution has been arranged, are few, the curvature of the curve becomes angular and interpolation is unreliable throughout

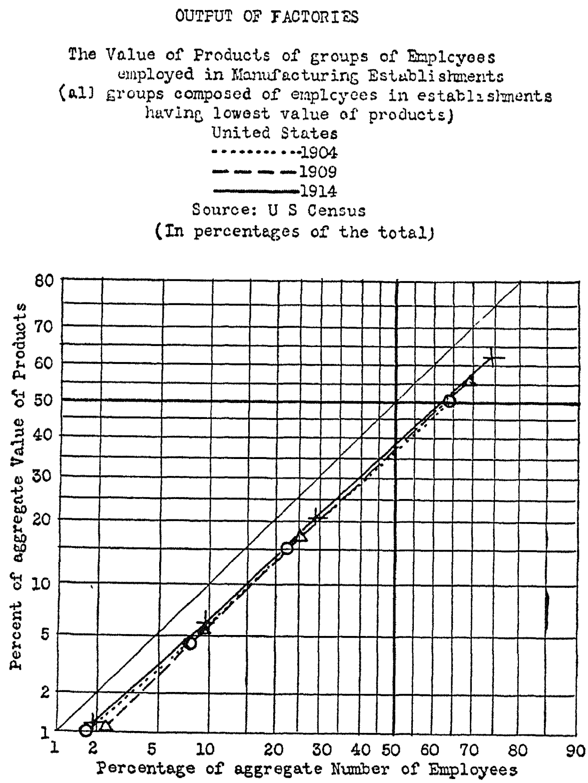


Fig. 383. The Double-Probabilities Projection Straightens Out the Lorenz Curves When of Normal Distributions.

the length of the curve. But when both axes are ruled on the probabilities scale, the tails become indefinitely long, the zero

PERSONAL INCOMES AND TAXES
Distribution of income and tax among tax-payers
United States
1919
(Source:- Collector of Internal Revenue)

INCOME CLASS	CUMULATIVE (Less-than) PERCENTAGES		
Over $1,000 or 2,000 and-	Returns	Income (net)	Tax (total)
Under $ 2,000	36.09	14.25	1.95
" 3,000	65.53	33.42	4.18
" 5,000	87.67	56.14	10.16
" 10,000	95.90	71.01	17.37
" 25,000	98.94	83.16	30.35
" 50,000	99.64	89.59	42.55
" 100,000	99.89	94.10	57.23
" 150,000	99.95	95.91	66.58
" 300,000	99.986	97.78	79.43
" 500,000	99.995	98.58	86.21
" 1,000,000	99.999	99.23	92.21
Total	100.00	100.00	100.000

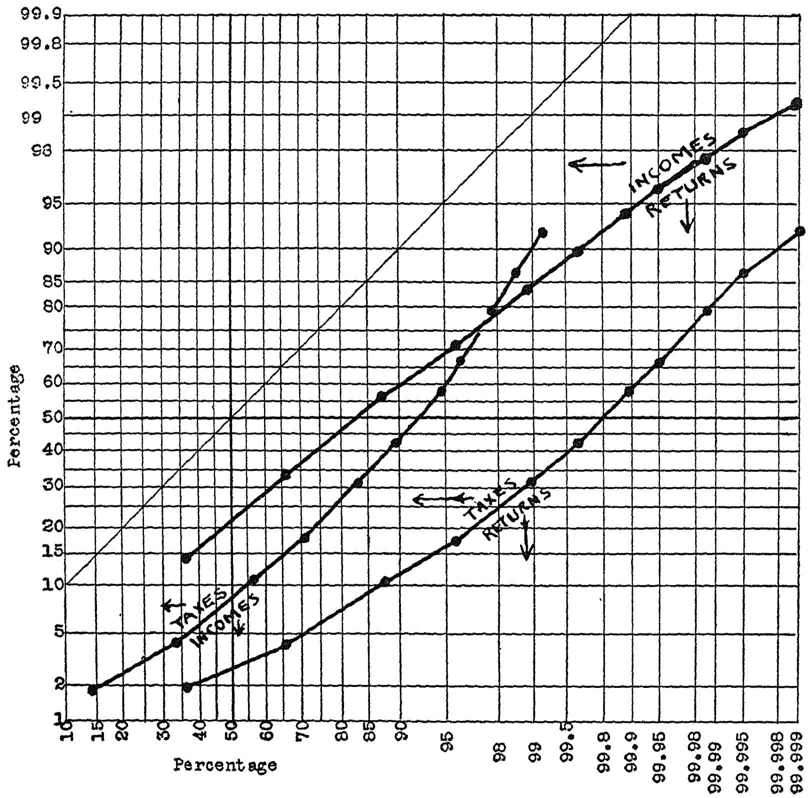


Fig. 384. Double-Probabilities Projection for Several Lorenz Curves. The chart shows that, for example, one-half of the tax-payers paid only 3% of the taxes, having only 22% of the income, so that one-half of the income of tax-payers yielded only 8% of the taxes. The curves are not straight because the tax-payers do not constitute the entire income-receiving population and form therefore an incomplete or truncated part of what is probably a normal distribution.

and hundred per cent points disappear, receding to infinite distances, and the curve, throughout its length, becomes very close to a straight line. When the distribution is normal, obviously the curve becomes a straight line, and hence the straight-line Lorenz curve on double-probabilities paper is a quick and useful indication that the distribution is normal. Obviously the accuracy of interpolation is improved by this straightening out of the Lorenz curve, as is also the facility for detailed comparison, such as through light-analysis, of several distributions so plotted.

The utility of the probabilities projection must be apparent to those who deal with frequency data. For analytical purposes, as a labor-saving and illuminating chart, it takes its place beside rate-of-change paper for historical data.

Chapter XLI

SHIFTED ZERO-POINTS

We have seen, thus far, three great types or kinds of scale projections, arithmetical or uniform, logarithmic or geometrical, and normal or probabilities. We have seen these combined in every way upon the two axes of the chart. All this has been done in the search for simplicity or regularity of behavior, and convenience or ease in interpolation. There remain still other projections of the chart-scales, which serve the same purposes and will be discussed in later chapters. And there are a few minor variations of the logarithmic projection which can well be discussed here.

We have seen that historical data can be plotted with either logarithmic or arithmetical vertical scales, but that it is not correct to use anything except an arithmetical scale on the horizontal axis. To this rule we may now note two exceptions. The first arises in the case of data with a definite origin point. Thus the pseudo-historical frequency series which involve time have already been put upon logarithmic x-axis scales together with other frequency series. From these frequency series in which time is the independent variable, it is but a short step to strictly historical series, in fact the distinction disappears here, the same series being called equally well a frequency or a historical one. But there is also a class of purely historical data, involving specific points of time, which can be placed upon a logarithmic x-axis. This is data, generally of a geological, or other scientific nature, covering very large periods of time, such as the age of the earth, and its important geological eras.

The second variation of historical series is extremely interesting, though of very limited application. It may be called the retrospective projection. If from any point of time we look backward over the years, we may notice that the more recent events stand out more clearly, and in more detail, while

the events of early years become more vague and their details lose importance. In business, this importance of the last years is recognized and business statistics therefore often contain full detail for the most recent period and only brief summaries of previous periods. In histories, the space devoted to ancient, medieval, and modern times, usually shows a similar compression of earlier times. If these witnesses are of any value, they testify that the importance of detailed data diminishes as its remoteness in point of time increases. And the chart-maker has therefore a legitimate object in devising a chart method to display the data in its proper detail or lack of detail..

Several methods have been tried to meet this charting need. By the silhouette bars a few facts of the past history are given in addition to the very latest figure. By the juxtaposition of two curves with a single y-axis scale but different x-axis scales, one, let us say, for years, the other for months, data for a recent period can be given in full detail, and that of a previous or the entire period in summarized form. But the inventive mind will seek still a better method, which will not have the rigidity of the last and in which the disappearance of detail will be gradual and so we arrive at the use of a logarithmic x-axis scale projection, reversed in its direction so as to compress the earlier periods of time upon the chart.

This retrospective logarithmic projection of the time scale has both intriguing advantages and baffling disadvantages. Upon its credit side we may observe that it presents precisely the degree of importance to data at various points along the line that we desired and has unlimited possibilities of extension backwards into remote antiquity without consuming space wastefully. If we are, in the year 1900, let us say, to look back over the centuries, we shall doubtless attach the same relative importance to the entire nineteenth century as we do to the seventeenth and eighteenth combined, the same relative importance to the last thousand years as to the two previous thousand years. Important exceptions occur, of course, but in the main, this proportion of weight of importance holds, else historians would not be justified in devoting their space to the different periods in these ratios. . The real nature of that phenomenon of change which we call the passage of time is still a profound mystery and it is an interesting speculation that it may in some occult way combine elements of progressive and regressive organic growth. But idle as this thought may be, the chart of

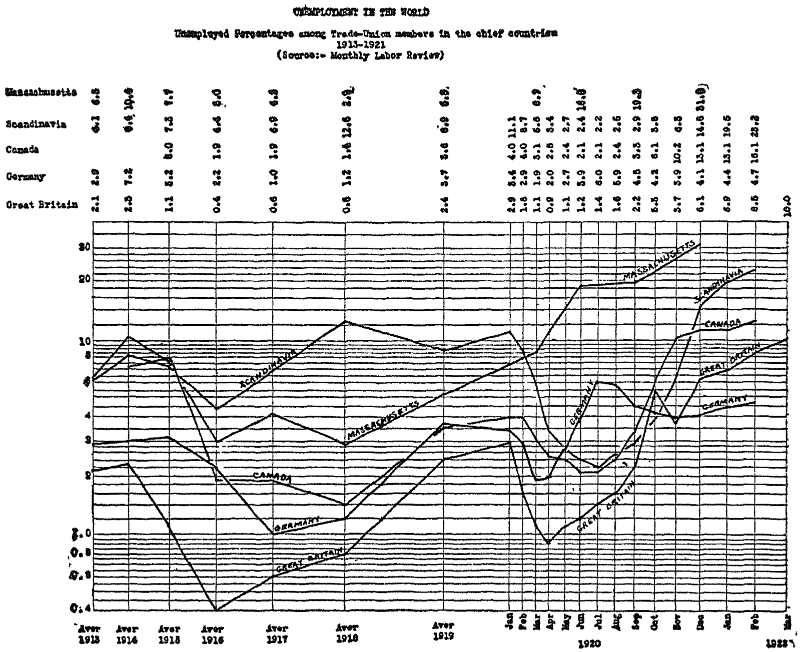


Fig. 385. A Historical Retrospect with Reversed Log Plotting for the Horizontal or Time Axis.

The purpose is to present recent developments in greater detail.

the geometrically retrospective historical curve has a certain value in presenting graphically a survey of the past in proper emphasis and detail, and affording a comprehensive picture not otherwise equalled. Owing to the fact that the significance of the slopes of the curve disappears in this chart, its usefulness in mathematical curve-analysis will always be limited, if not doubtful. But when the gun-shot method of plotting be employed (that is, isolated points be plotted) its success is marked. It is not improbable that in time all school histories will be illustrated with diagrams in which events will be entered at their proper positions upon such a scale.

The chief disadvantage of the logarithmic retrospect chart lies in the fact that before entering the time-figures upon the scale, these figures must be computed back from some origin point, either in the present or in the future. The choice of origin-point for our backward count of the months or years directly effects the degree of expansion which the most recent periods of time will undergo on the chart-scale. If we take the

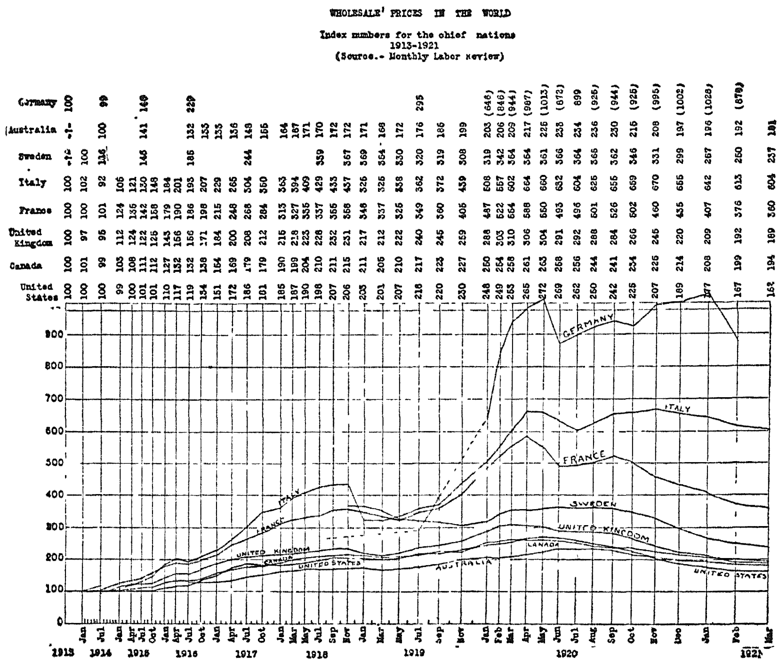


Fig. 386. Another Example of the Same.

immediate present, then last month's figures will be plotted at unity, as being one month back, the previous month's data will be entered at the logarithm of two, on the left side of unity, and so on backward along the axis of the chart. But where will we plot next month's figures when they appear? We cannot plot them at the logarithm of zero. In other words such a retrospective chart would have to be redrawn each month. The best way to avoid this is to assume an origin point of time some distance in the future. Although this destroys the significant relation of the more recent events, so far as that relation was desirable, yet it enables us to bring the chart up to date for some time to come. But the student will now see that comparisons cannot be made from one such chart to another unless common "view-points" have been used in both. The disadvantage is so serious as to make the chart useless for research purposes, but it is still worthy of notice for general records and popular presentation.[1]

[1] For short periods of time, a "squares" or other powers projection will serve equally well and can be continually added to.

The use of altered logarithmic scales for the frequency series is a step which requires much more technical mathematical justification. We seek here to bring not detail nor convenience, but symmetry and regularity, to the curve. Experiment will show that a great many asymmetrical distributions can be made to approach symmetry by assuming false origins and correspondingly altering the logarithmic scale projection. That this is necessary for data in which the original zero is a false or arbitrary one, having a real positive value, has already been pointed out. But when cause for shifting the zero does not clearly exist in the very nature of the data, the student should be slow to alter it, for even the most excellent symmetry which may be induced thereby may be utterly lacking in significance. This applies not only to shifting of zero points, but also, of course, to reversing of directions from the zero, a step which is closely allied in that it amounts to giving to some value above the maximum of the range an assumed value of zero and treating the resulting negative values in the range as if they were positive.

There is open to us another alteration of the x-axis scale projection which is not of any value for historical curves, namely, the anti-logarithmic projection. For just as we have been able to plot the logarithms of our scale-figures, so, too, we can plot their anti-logarithms. The device has results similar to the retrospective projection, in that it expands the larger numbers and compresses the smaller ones. Here, too, caution must be used in attaching significance to the results. But the mathematical interpretation of this projection, though still technical, is much simpler. There are in fact certain classes of measuring units which are definitely of a geometric nature and not arithmetical. It would be useless to plot these upon logarithmically projected scales, for the numbers themselves are already logarithms, and the log scale would really give a log-log projection. Thus stars are classified as of various magnitudes, each magnitude being two and a half times as bright as the next. Musical pitch is measured in tones and octaves, each octave having twice the wave-frequency of the preceding octave. In these cases if we seek an alternative projection we must obviously use the anti-logarithms of the magnitudes and tones. Both the anti-log and the log-log projections may occasionally be useful upon either axis of the chart.

Chapter XLII

CURVE-FITTING

The reader has now seen a variety of ways by which symmetry or regularity can be brought to the curves of historical and frequency data. The reason, as he has seen, behind this quest for simplicity, for straight lines, for parallel or mirroring curves, and the like, lies in the ease with which generalization can proceed from such forms, and the degree of confidence with which we may accept the data as reliable samplings and their curves as significant pictures. When we see the curve of the country's population mounting higher so steadily that it forms one single straight line, we know, without performing any mathematical exercises at all, that the population grows at a constant rate, and we are thankful to the logarithmic projection of the chart which has yielded this simplicity. When we see the curve of our advertising appropriation paralleling, on rate-of-change paper, the curve of our gross sales, we know without any computing that the same fraction of the dollar has gone into advertising every year. When we see the cumulated curve of the nation's population as divided into cities of various sizes, forming a straight line upon the probabilities paper, we know that without appreciable error or effort we can by interpolation find the number of persons inhabiting communities of any particular size, whether or not the Census has mentioned communities of such size. And in every case, the same causes assure us some degree of confidence in the reliability of our observations as fair samplings, when we have put partial data.

The question of the reliability of data is properly a statistical one, for which the student should consult the statistical authorities.[1] It arises in the collection of data, before charting has begun, and only recurs again when the near approach of a charted curve toward regularity and simplicity raises the sug-

[1] See particularly Bowley, p. 178.

gestion that the deviations of the actual curve-line from the desired simplicity of form are due to errors of data. Thus if the ogive of the distribution of incomes is very near to a straight line on logarithmic paper, it is but natural that the theory should arise, as at least a tentative explanation, that the deviations are due to omissions in tax collection, evasion in income reporting, or in some cases chance variations due to few observations.

And it often happens that in the desire to justify a theoretical simplicity, we are too ready to excuse deviations from it as errors in the data or chance variations. A straight line may fit so closely to the data that we feel sure that it, instead of the observed curve, represents the truth. Thus the entirety of Pareto's law of incomes is a result of adopting the fitted straight-line rather than the actual income curve (ogive). Since the income curve is truncated at its lower end owing to lack of information on incomes below the tax limits, that law is based upon insufficient data. And recent investigations lead to the belief that the true ogive of incomes is not a straight line upon logarithmic paper, but upon logarithmic probabilities paper.[2] Slight deviations from a straight line are not conclusive evidence of errors in the data and the student should be extremely careful in drawing hasty conclusions from a close approximation to a straight line upon the special projections which have been described.

There is, therefore, always a danger that the close approximation of a given curve to a straight-line, or other simple theoretical curve, is fallacious and deceptive. This caution cannot be too strongly emphasized. It attaches *prima facie* to all attempts to fit theoretical curves to actual ones and casts upon him who would fit such curves the burden of proof. The presumption is that the deviations of the given curve, from the fitted one, are significant. The removal of this presumption may call for all the analytical powers of the statistician, but we should always start with the presumption, and never lightly abandon it.

There may be many reasons why the deviations are insignificant. Some of these may be found in the particular circumstances surrounding the collection of the data, such as bias on the part of the investigators, or difficulties of observa-

[2] See National Bureau of Economic Research, *Income in the United States*.

tion. But, whatever else may be found, there is likely always to be one cause for the lack of a perfect fit, in what are commonly called "chance variations." These are more marked in small samplings than in large ones. A definite mathematical law for the probability of this occurrence can be found in books on the subject. The reader who recalls the normal curve of error will understand the inevitability of such chance variations. And, needless to say, when we can safely consider the deviations of a given curve from a fitted theoretical one to be due to chance variations, these deviations lose all significance and we may safely proceed with the fitting.

The theoretical curve which we propose to fit to a given curve may have any shape. The simple linear curve or straight-line upon plain paper with arithmetically-projected scales, is merely the simplest of these. And the reader will remember that, in the discussion of cycles in historical data, the fitted straight-line was called the "secular trend." It is a very crude secular trend, convenient, but in most cases not precise. The reader has since seen that, for most economic data, a straight-line upon the rate-of-change (or semi-log) paper would be more accurate. Still other "trends" and straight-lines will be described in later chapters. For frequency data, the fitted curve is generally the normal curve, or its equivalent straight-line upon probabilities projections. Of the significance and appropriateness of these straight-lines, in each case, the reader has already a general understanding, and the mathematics of these and other straight-lines will be discussed later.

While so much attention is being given to the various charting methods by which theoretical curves are reduced to straight lines and by which actual curves are more easily compared with theoretical ones, it would seem well to mention briefly the mechanics of fitting. This problem arises after the particular theoretical curve to be fitted has been chosen. Let us assume that it is a straight line upon one of the chart-forms already described, such as the semi-logarithmic or rate-of-change paper, or the probabilities paper, or even the plain uniform paper with arithmetically projected scales. The problem of fitting is virtually the same in all cases.

Whenever a straight-line (in the example we have taken) is to be fitted to a given curve, and that curve does not form in itself a perfectly straight line, it is obvious that the straight

line may lie in an infinite number of slightly different positions and still fit very closely to the given line. The problem is, therefore, to find the particular straight line (or other theoretical curve of the selected type) which gives the best of all the possible fits. If we call the deviations of the given curve from the fitted one, its "residuals," the problem is, broadly speaking, to find the fitted curve which makes the total of these residuals a minimum (that is, the least possible sum for the given curve).[3]

There are three outstanding methods which have been developed for determining the best fitted straight-line. These may be called the graphical method of selected points, the method of averages, and the method of least squares.[4] Of these, the first is the simplest; the last, the most accurate; and the second, the most satisfactory because both fairly simple and fairly accurate. The graphic method of selected points is nothing more than laying a transparent straight edge or tightly drawn piece of thread over the curve and adjusting its position until an equal number of points appear on both sides of the straight line and the fit appears optically most satisfactory. The other two methods are mathematical processes, for which the reader will have to consult the proper statistical authorities.

But the first of these mathematical processes for determining the position of the fitted straight line, namely the method of averages, is also capable of a graphic solution. If you join the first and second plotted points and plot a new point midway on their joining line, this new point will represent their average. If you repeat the process with the third and fourth, the fifth and sixth, and so on with each successive pair, you can reduce the whole curve to a slightly shorter curve with only half the number of plotted points all of which are averages. On the new curve fresh averages can be plotted, this time representing averages of averages, or averages for four points on the original curve. After repeating this operation a sufficient number of times, you can reduce the longest curve to a series of two average points, through which a fitted straight line can be projected.

[3] Because the algebraic sum of the differences from the mean is always zero, statisticians often use the squares of these deviations. Strictly speaking, therefore, the problem is to find the fitted curve which makes the total of the squares of the residuals a minimum. If the residuals alone, instead of their squares, be used, we must seek to make the arithmetical sum (that is, the sum of the residuals, disregarding their signs) a minimum.

[4] *Cf.* Merriman's *Method of Least Squares* or Bartlett's *Method of Least Squares.*

For many purposes it is sufficient to fit curves by inspection, just as it is sufficient to correlate them in this way. The graphic analysis, made more precise by "light analysis," is a tremendous labor-saver, and may serve at least in the preliminary stage of the study, at least.

In fact, curve-fitting is but a variation of correlation, being merely the determining of the theoretical curve which best fits or correlates with the given curve. And the considerations affecting correlation likewise govern curve-fitting. For precise purposes, the mathematical method of least squares should be employed and a mathematical coefficient of correlation and probable error be computed to measure the success of the fit.

Chapter XLIII

SPECIALLY PROJECTED SCALES

We are about to embark upon an orgy of distortions, modifications, and special projections of the scales for curve-charts, all of them being designed graphically to facilitate the study of particular data. The general principles and the more useful forms for the non-mathematical reader will be set forth in the present chapter. In the succeeding chapter, the mathematics of all special projections will be discussed, from which any particular projection can be designed. The present chapter will suffice for most.

It will by this time have occurred to the reader that the scale figures or calibrations may be plotted or graduated at any points along the axis of the charts which we desire, and can therefore be made to express any function of the variable plotted thereon. Thus the logarithmic projection is merely one in which, if we can designate by X the scale-figures or calibrations and by x the actual distances at which these are placed along the axis, then in the logarithmic projection $x = \log X$ (and $X = \text{anti-log } x$ and $10^x = X$). From this it is no difficulty to proceed to the scale projection of other functions of the variables. A very simple example of this would be the expression of reciprocals by the scale $x = \frac{1}{X}$ (or $X = \frac{1}{x}$). Obviously such a scale along an axis would straighten out all curves in which one variable varied with the reciprocal of the other, and if both axes be plotted on such scales, then curves would straighten out when the reciprocals of both variables vary together.

Perhaps to the economist the most interesting application of special scales lies in a recently discovered use of what may be called a "square-root projection" for certain historical data. The speculation which leads to the use of this projection is founded upon the analogy of familiar physical laws governing

the intensity of light, the flight of falling bodies, and the like, in which one set of values varies (directly or inversely) as the square of another. In the case of light, as every one knows, its intensity varies inversely with the square of the distance from its source and the area of the cross section of a beam of light varies directly with the square of the distance. Falling bodies travel in each unit of time over a distance proportional to the square of the number of units of time they have been falling, and their velocity therefore varies with the square of the length of time elapsed since leaving a position of rest. The idea suggests itself that certain economic phenomena may closely parallel in their growth the growth of such natural phenomena. It would be obvious, of course, that this method of analysis could only be applied to phenomena which are free of elements of organic growth or other factors, or in which corrections can be made for such elements and factors.[1]

Other powers and roots may equally well be the basis of the special projection. These have not, however, as yet become important to the economist; they are chiefly useful in engineering and the natural sciences, where formulae and equations are found of every type, and degree. The conic sections, the circle, ellipse, hyperbola and parabola are all important to the scientist, while the economist does not need to go beyond the parabola and the hyperbola when he leaves the straight-line. Indeed the statistician either in business

---

[1] The Gompertz curve, as it is sometimes called, to which much economic data fits, is not unlike the curve which straightens out on a square-root projection, when the curve has been plotted upon a logarithmic projection instead. In other words, the square-root projection suggests itself for all economic data in which the curve, like an ogive, seems to have a "die-away" approach to a maximum when plotted on log-paper, as if reaching a saturation point.

The Gompertz curve is, however, probably nearer to the typical behavior of economic phenomena through their initial stages, from discovery and through experimentation and installation, to the final stage of "saturation" in which maintenance and upkeep constitute the chief requisites. This curve has not yet been made the subject of a special projection. Its formula is

$$y = ab^{c^x}, \text{ or } \log y = \log a + c^x \log b, \text{ or } \operatorname{loglog}\left(\frac{y}{a}\right) = x \log c + \operatorname{loglog} b$$

indicating that a loglog y-scale with shifted zeros and an arithmetical x-scale will straighten the curve, when the value of the constant a has been determined. For a recent excellent discussion of this curve and the methods for determining the constants, the reader should see Prescott, Raymond B., *Law of Growth in Forecasting Demand*, in the Journal of the American Statistical Association, December, 1922, pp. 471–479. See also Running, Theodore R., *Empirical Formulas*, John Wiley & Sons, New York, 1917, pp. 29–33.

THE WORLD'S COMMERCE
Estimated commerce and commercial equipment specified
World, 1800-1919
(Source:- U. S. Statistical Abstract)

	1800	1820	30	40	50	60	70	80	90	1900	06	10	13	15	17	1919
Commerce ($-,000,000)	1,479	1,659	1,981	2,789	4,049	7,246	10,663	14,761	17,519	20,105	27,418	33,634	40,420	31,502	52,781	76,513
Sailing vessels (-,000 tons)	4,026	5,814	7,100	9,012	11,470	14,890	12,900	14,400	9,168	6,674	5,809	4,624	5,891	5,533	4,425	6,046
Steam vessels (-,000 tons)		20	111	368	864	1,710	3,040	5,880	8,295	15,857	19,713	22,046	26,517	28,160	25,580	28,177
Railways (-,000 miles)			0.3	5.4	24	67	140	225	390	500	564	657	690	718	730	751
Telegraph (-,000 miles)					5	100	281	440	788	1,180	1,200	1,307	1,462	1,526	1,568	1,577
Cables (-,000 miles)					0.025	1.5	15.	49	152	200	270	291	330	336		

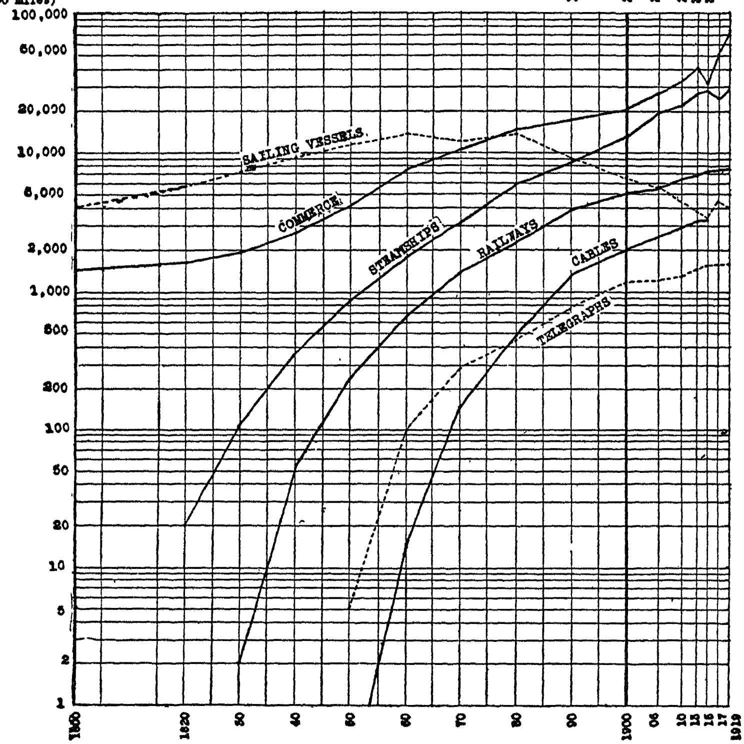


Fig. 387. The Four Lower Curves Fail to Straighten Out on Logarithmic Vertical Scale.

or economics rarely works with such precise and inflexible data as the scientist, and cannot so often attempt precise mathematical generalization, in the shape of formulas and equations.

THE WORLD'S COMMERCIAL EQUIPMENT
Estimated Railways, Steamships, Cables and Telegraph
World, 1820-1919
(Source:- U. S. Statistical Abstract)

	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1915	1919
Cables (Thousand miles)				0.025	1.5	15	49	132	200	291	335	
Steam Vessels (Thousand tons)	20	111	368	864	1,710	3,040	5,880	8,295	13,857	22,046		28,177
Telegraph (Thousand miles)				5	100	281	440	768	1,180	1,307		1,577
Railways (Thousand miles)		0.2	5.4	24	67	140	225	390	500	637		731

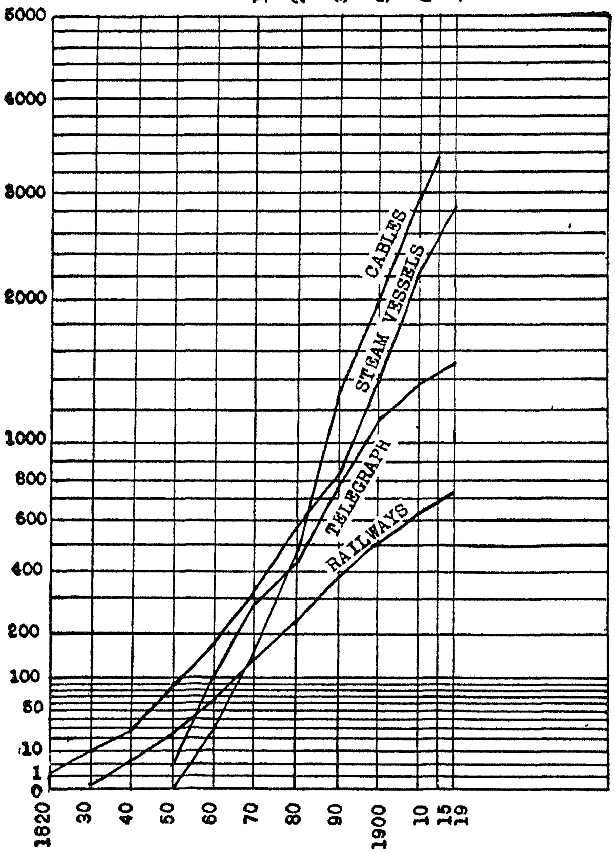


Fig. 388. The Square-root Projection of the Vertical Scale Brings Much Greater Regularity to the Curves of the Preceding Chart.

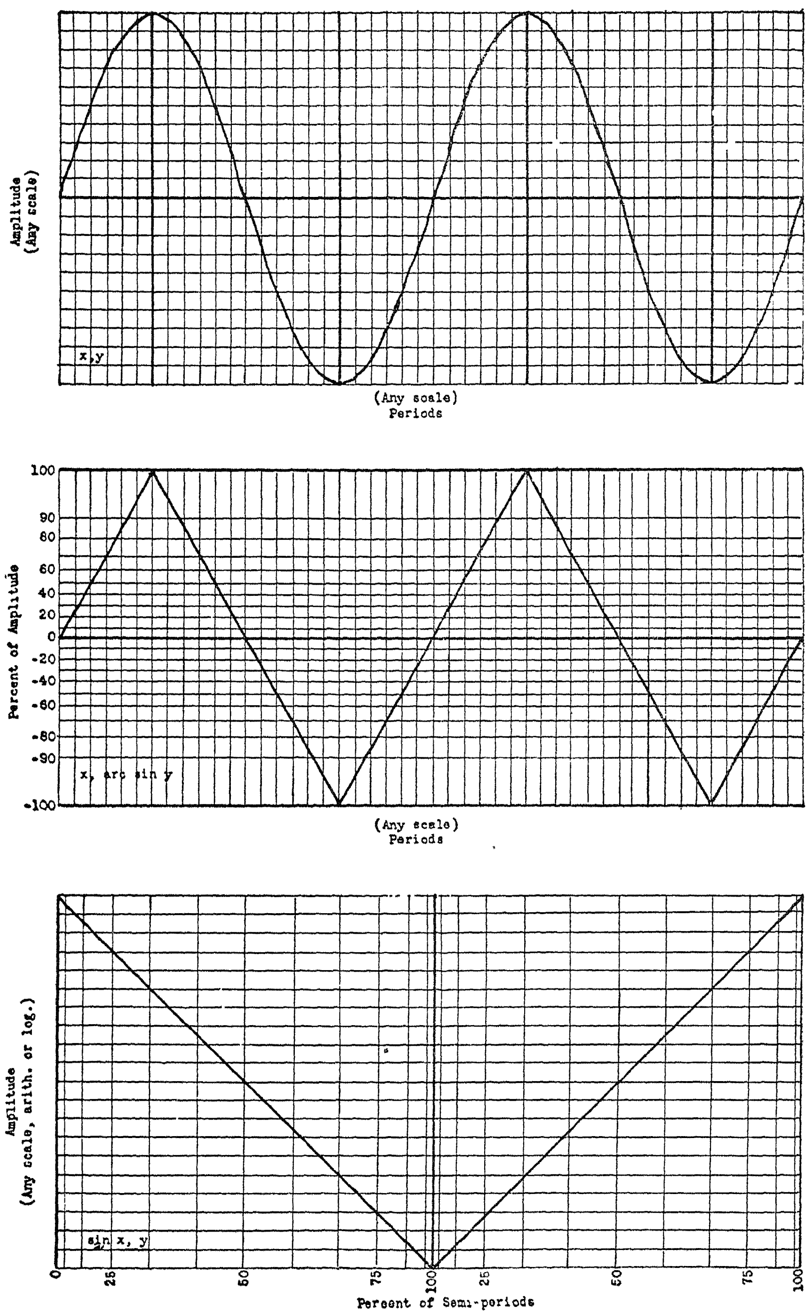


Fig. 389. The Two Ways of Straightening Out Semi-cycles of a Sine Curve.

It is a rule of general application that any curve can be easily made straight if it does not undulate, that is, has neither peaks nor valleys. The process of straightening requires no more than the projection of the y-axis scale, that is, the scale of the independent variable, in such a way as to make its calibrations record the values of the ordinates at equal intervals along the x-axis scale of the arithmetically projected chart. A more detailed specially projected scale is given by the following steps: 1, Divide either axis of the arithmetically projected chart into uniform parts or intervals; 2, From the points so obtained erect ordinates or abscissae to the curve, and from their intersections with the curve project abscissae or ordinates to the other axis, and read the values thereon; 3, Lay off these values along the first axis but calibrate them with the original values. The object of this is always the same, namely, to introduce in the altered scale of the variable those inequalities and irregularities which will exactly counterbalance the irregularities of the curve and smooth it out into a straight line.

It will quickly occur to the student that such specially projected scales can even be made for some undulate or periodic curves, to reduce them to regularity and uniformity of undulations. By undulate curves we mean curves with peaks or valleys or both. Thus a sine curve can be laid out on such specially projected scales so that it forms a succession either of semi-circles or of angles and straight lines. It is in such work, however, most obvious that the peaks should be correctly positioned at the maximum ordinates or abscissae, the valleys at the minimum ordinates or abscissae and where cycles are different in amplitude or phase, that they be converted to common levels or intervals. The intricacy of this work is considerable, its usefulness highly specialized and the economist, sociologist, or business statistician will have little occasion for it.

The subject of special projections and their use for the analysis of curves is still in an elementary stage and little can be dogmatically stated about it. Collections of the typical

---

Note to Fig. 389

The first requires the conversion of all amplitudes into percentages of the nodes, and permits the use of variable phases and periods. The second requires the conversion of all cycles to common units of π or percentages thereof, and permits any amplitude readings.

hyperbolic, parabolic, and other curves have been published, intended as guides to the engineer to assist him in the recognition of the nature of given curves. Such collections familiarize the student with the curves of his equations, but are restricted in their usefulness in the reverse process of equating

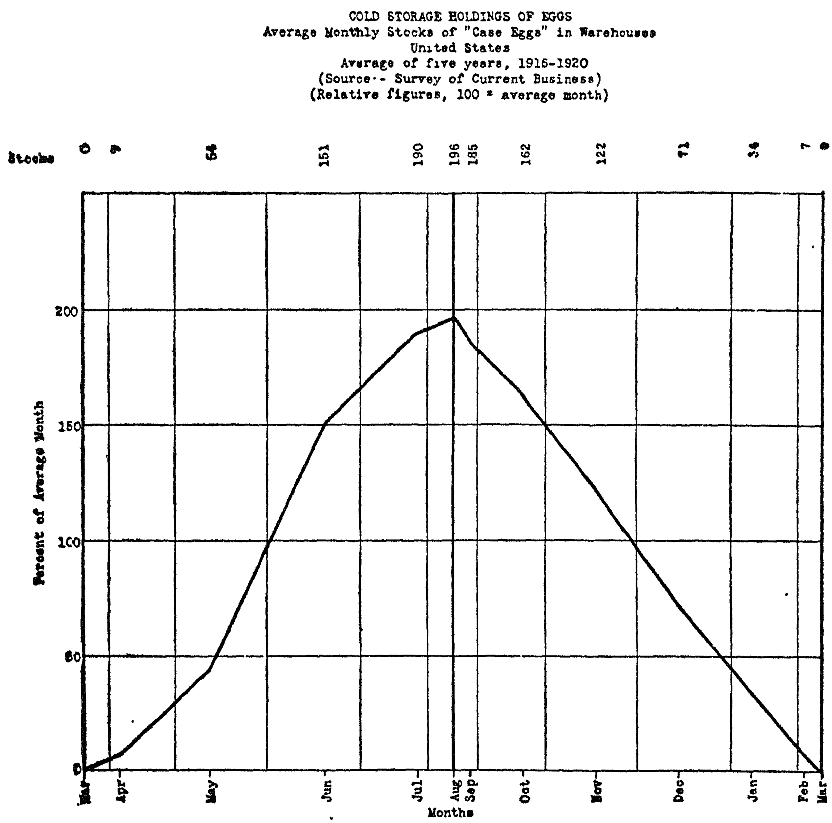


Fig. 390. Showing How Closely the Cycles of One Set of Periodic Economic Data Approach a Sine Curve Wave.

curves. The trouble arises in the fact that slight changes either in the scale of the chart or the constants in the equation produce great changes in the appearance of the curve. It is therefore impossible to prepare complete catalogs of curves classified by their shapes.

It is probable that in the development of special scale projections which straighten or regularize the curves of certain equations, the greatest advance in the science of curve equat-

ing will be made. For very often these special scales are free of the limitations of the fitted curve, neither scale alterations nor changes of constants affecting the regularity of the curve upon the chart with the correct special scale. This is a field of graphics as yet little explored, its possibilities are just opening up to us, and only time and continued usage will determine what forms are valuable and the limits of their value. It is clear, however, that the probabilities projection is not the last invention of its kind nor yet is the square or powers projection.

Chapter XLIV

FORMULAE FOR CURVES

While it is a great step forward in the analysis of data to have plotted the curve and be able to visualize the behavior of the phenomenon, yet the mathematician often seeks to take a further step, and formulate from the curve a law for the data, by which its behavior will be precisely described in a single mathematical sentence. This mathematical sentence is known as an "equation." And it is the object of the equation to give us a general description of the data under all circumstances and times for which the equation is prepared. The process of describing the behavior of a curve with this mathematical precision, is called writing an equation to the curve.

For a great many curves the description is extremely easy to formulate. If you see a straight line on amount-of-change paper, in which the two scales are alike and the straight line slopes at an angle of 45° from the origin of the chart, you will say at once that the y-values which the line passes through are equal to the x-values, or that the y-values are increasing equally with the x-values; and you would express this description mathematically in the sentence (or equation) $y=x$, that is, that the value of y for any x is the same as the x itself and whatever the x-value is, that also will the y-value be. Let us vary the case a bit. If the y-scale be twice as great as the x-scale, that is, each unit on the y-scale equal to two units on the x-scale (the line still sloping at 45 degrees), it would not take you long to determine that the formula or equation for the straight-line curve is $y=2\,x$. Again, let us suppose that instead of passing through the y-axis at the origin, the curve passes through it at the value of 3. Now this adds 3 to each value of y throughout the length of the curve and so you will quickly write the formula as $y=2\,x+3$.

Or let us take a descending curve which on a chart with equal scales on each axis describes a 45° downward slope from,

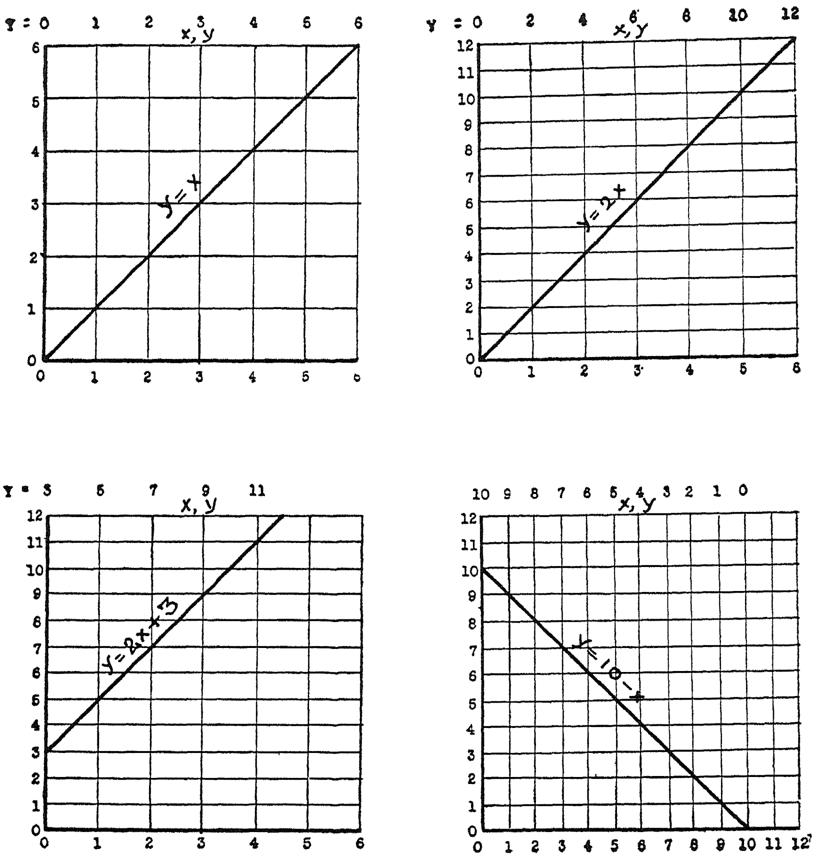


Fig. 391. The Linear Equation, $y = ax + c$.

EXAMPLES:

$y=x$ $\qquad$ $y=2x$

$y=2x+3$ $\qquad$ $y=10-x$

The value of c can be read at the intersection of the curve with the y-axis; in other words c is the ordinate of the curve when $x=o$. The sign (+or−) of a is shown by the upward or downward direction of the curve; its value can be found from the phrase, $a=y-c$, when $x=1$.

(Note.—The small letters immediately above each diagram in this chapter indicate the functions plotted to form the curves.)

let us say, the point of 10 on the y-axis. A little study will show you that as the curve descends, the values of y diminish from their original value by the amount of the corresponding x-values. This condition can be expressed mathematically by the sentence, $y=10-x$. All of these cases are simple and we can observe that in all of them the curve forms a straight line. So we may hazard a guess that whenever we meet a straight line

upon amount-of-change paper, we can write an equation to it which will be a simple equation or an equation of the first order, that is, both the unknowns, y and x, will be found in their first powers. The general formula for the straight line on plain co-ordinates is sometimes written as $y=ax+c$, in which both a and c are constants for the particular line. (The values of these constants were seen in the examples just given to have been, for "a," 1, 2, 2, and -1; and for "c," 0, 0, 3, and 10.)

Remembering that the logarithmic projection substitutes the processes of multiplication and division for the processes of addition and subtraction, we may further generalize that a straight line upon logarithmic paper will have the general formula of $\log y = b \log x + d$, or $\log y = b \log x + \log a$, which is the same as saying $y=ax^b$. And a little experimentation will show you that every equation of this form (a and b being constants) will appear as a straight line upon a logarithmic chart. Here again we find a simple formulary relation which it is convenient to determine. For an equation so simple as $y = ax+c$ or $y=ax^b$ is distinctly more convenient to remember and apply than the plotted curve itself. Two familiar examples of this are to be found in the computing of interest, the first being the formula for simple interest plus principle and the second for compound interest plus principle.

The special scales which have been discussed in the previous chapter are of course designed expressly to whip into straight line formation the recalcitrant and unwilling curve, and when they succeed, or even very nearly succeed, greatly simplify the writing of equations. But as has been pointed out, they can only be used with care, since some curves, or short portions of curves, will behave similarly upon several projections, and the approach to a straight line upon one scale projections does not always indicate that the formula for that scale is the best, or even a correct, formula for the curve. This danger has already been mentioned and illustrated.

A little study of the various special scale projections will show that they are all outgrowths of the simple linear equation of the straight line upon uniform or arithmetically projected scales. The object of the special projection in each case is to so graduate the values of the scale that they absorb all the powers of the variables, leaving to be plotted the remainder of the equation, in which the variables occur in the first powers only and which therefore form straight line curves on the

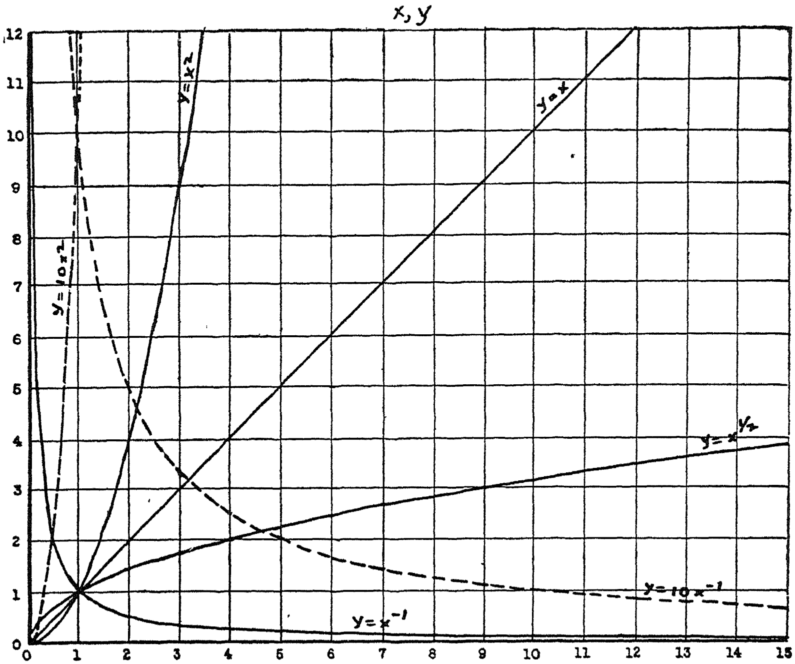
chart. Lipka enumerates eleven typical equations whose curves straighten out upon the charts with the scales and we shall briefly repeat this list, that the student who has found a combination of scales which makes his curve straight may quickly find the equation best describing his data.[1]

For the straight line upon uniform (that is, arithmetically projected) scales, we have an equation in the first degree, called, from its form, the linear equation. Its type is $y=ax+c$, in which a and c are constants whose values can be easily found, c being the intersect point of the curve upon the y-axis and a being the tangent of the angle of the curve upon the x-axis, easily computed from any observation after c is known. For all curves which pass through the origin, the equation is reduced to $y=ax$, since c has disappeared.

For the straight line upon log paper, both scales being logarithmically projected, we have, as we have seen, the equation $\log y=b \log x+d$, which is but another way of saying $y=ax^b$ in which a, b, and d are constants, d being the logarithm of a, and both d and b are as easily found as c and a above. Drawn upon arithmetical paper, the curve is, of course, not a straight line, but becomes a simple parabola or hyperbola. It is a hyperbola, that is, in this case, a falling curve, if b is negative; if b is positive, the curve is a parabola, that is, in this case, a rising curve, and approaches the vertical as it increases if b is greater than unity and approaches the horizontal if b is less than unity. In the one case where b is unity, the curve straightens out (on arithmetical paper) since here b can be omitted from the equation and the latter becomes $y=ax$. Thus we see that the equation $y=ax$ will be a straight line upon either the arithmetical or the logarithmic projections. This is the same as saying that if a straight line curve on arithmetical paper pass through the origin, it will also be a straight line upon logarithmic paper.

We have in the previous chapter mentioned the shifting of the zero point upon a logarithmic projection, that is, the recalibration of the scale after it has been graduated (plotted.)

[1] The remainder of this chapter is largely and very inadequately drawn from Professor Lipka's excellent book, *Graphical and Mechanical Computation*, John Wiley & Sons, 1918. This has been done not to substitute the present volume in any way for that treatise; it has rather been the writer's purpose to draw attention to the extraordinary possibilities opened up by Lipka's work, and to direct readers to it. The volume is indispensable to the student, and to the technician it will almost certainly open up a new world of research, arming him with invaluable implements.



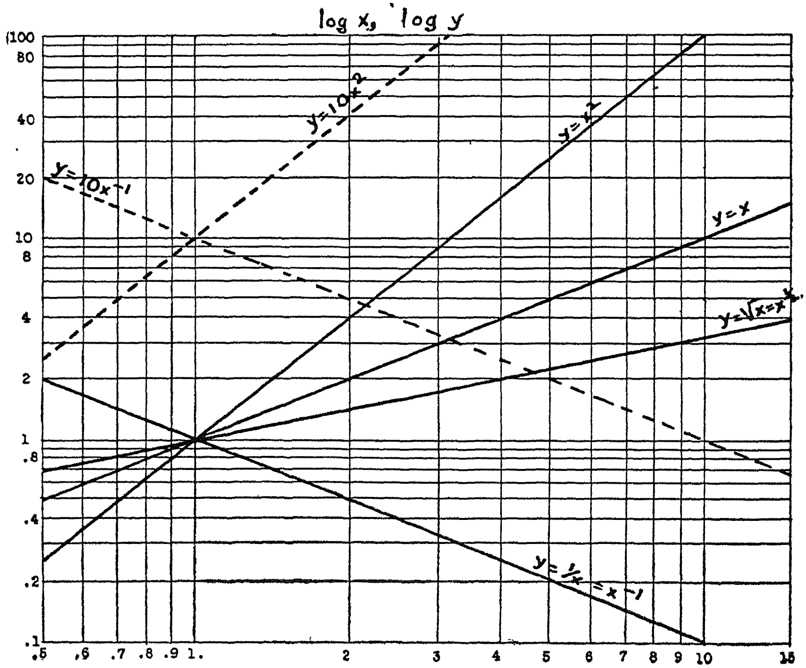


Fig. 392. The Curve of $y = ax^b$ or $\log y = \log a + b \log x$.
See footnote on opposite page.

The scale then represents, of course, log $(y-c)$ where c is the constant which has been added to the plotted values to give the calibrated scale-figures. When such shifting of the scale has straightened out a curve upon logarithmic paper, the curve has, of course, the equation log $(y-c)=b$ log $x+d$ (instead of the immediately foregoing log $y=b$ log $x+d$). From this new equation we derive $y-c=ax^b$ and so $y=ax^b+c$. Thus we see that the shifting of zero-points on the log scale is but an adjustment which makes c disappear and gives the straight line on log paper. On log paper without shifted zeros, the curve is parabolic, concave to the x-axis if c is positive, convex if it is negative.[2] When c is zero it disappears from the equation and the latter becomes $y=ax^b$, an equation already described, having the straight line form on unshifted scales of log paper. And when b becomes unity it disappears from the equation leaving an equation of the first type, forming a straight line on arithmetical paper.

The three types of equations are closely related, all having the general form $y=ax^b+c$, in which b is unity for the first type and any number for the others and c is zero for the second type and any number for the others. The third is distinct from the first and second in that it contains not one or two, but three constants, and hence requires calculation (when we are seeking to find the proper shifted scale) for the third con-

[2] The value of c can be computed by taking any three items in the data (or points along the curve plotted experimentally on uniform paper) such that their x-values form a geometric series, *i.e.*, $x_1:x_2::x_2:x_3$. Then

$$x_2=\sqrt{x_1x_3}$$

and

$$ax_2b=\sqrt{ax_1^b\ ax_3^b}$$

Hence

$$y_2-c=\sqrt{(y_1-c)\ (y_3-c)}$$

and

$$c=\frac{y_1y_3-y_2^2}{y_1+y_3-2y_2}$$

So we must observe the ordinates, y_1, y_2, y_3 at these points and substitute them in the last equation to get the value of c.

SCALE PROJECTIONS OF FIG. 392

Arithmetical

Logarithmic

The value of a is shown by the ordinate of the curve when x=1 (i.e., log $x=0$). The curve is hyperbolic when b is negative and parabolic when it is positive. (When $b=0$, the curve is a straight line parallel to the x-axis. If $b=1$, the curve is also straight upon the arithmetical projection, its equation, $y=ax$ being linear (see the curve $y=x$). The value of b can be found from the phrase, $b=\log y-\log a$, when $x=10$ (i.e., log $x=1$).

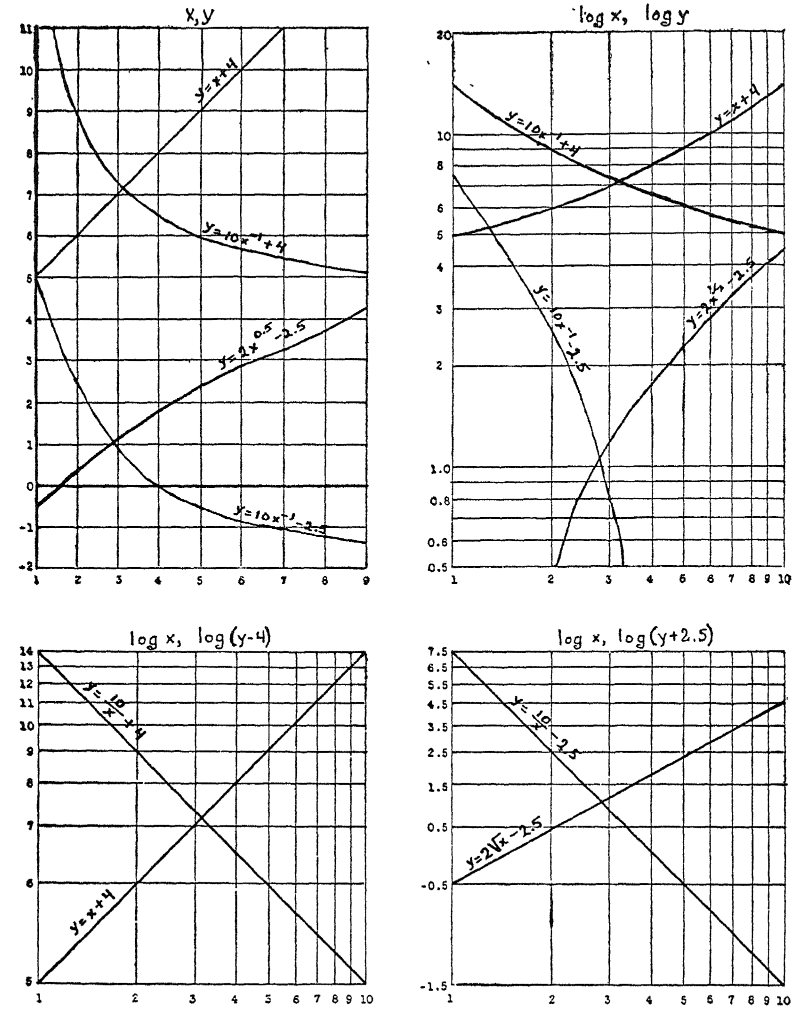


Fig. 393. The Curve of $y=a^b+c$ or, $\log (y-c)=\log a+b \log x$.

EXAMPLES:

SCALE PROJECTIONS:

Arithmetical *Logarithmic*

Logarithmic with shifted zeros

The scales of log (y-c) in the two lower diagrams are logarithmic projections with the scale-figures altered by the value of c; when c is negative the scale of log (y-c) will include the value of zero. The value of c must be known before the scale can be so altered. Four curves such that they straighten out on these shifted scales, are shown, two curves for each scale, one ascending and the other descending. These four curves are also shown on other projections, showing their various shapes. The value of a can be found by the phrase, $a=y-c$ when $x=1$ (i.e., $\log x=0$). The value of b can be found from the phrase, $b=\log (y-c) - \log a$ when $x=10$ (i.e., $\log x=1$).

stant, c. For the third constant is not readily capable of graphic solution, save on arithmetic paper; it must always be known or mathematically calculated before a proper altered scale can be found. Of course when we have by experiment found a satisfactory scale it amounts to a trial and error method of solution.

We also noted in the previous chapter the projection of powers of variables along the scale, the use of the square-root projection being illustrated. In these scales the values of x have been entered as calibrations or scale-figures at points which were plotted or graduated for the values of x^k, "k" being the known exponent of the power. A curve therefore which straightens out upon this when the y-axis is uniform (arithmetical), has the formula, $y=ax^k+c$, in which the constants, a and c, are found as before. This equation is of the same general type as the foregoing, $y=ax^k+c$, its only difference being that b is a known, not an unknown constant and is called k therefore. Here b or k plays the role of the third constant, being known. And when b is a small positive integral, such as 2, this scale projection affords a simple means of straightening the curve, but unlike c, b cannot be easily calculated when it is unknown. The method therefor is limited to the use of curves in which b is known, and is valuable in such cases when b is a small integral. It is much easier, for example, to prepare a squares projection than to shift the zero-point on a log scale.

The squares projection is an example of the powers projection in which the exponent of the power is a positive integer. If on the other hand, the exponent be fractional, we have an inverse power or root. An example of this would be a square-root projection. Or the exponent may be negative. The simplest instance of this is the reciprocal projection, for the reciprocal of a number is its -1 power. When a curve straightens out on a chart one axis of which is reciprocally projected (the other arithmetically) its formula is obviously $y=\frac{a}{x}+c$, a modification of the general type formula $y=ax^k+c$ (in which $k= -1$) or $y=ax^b+c$ (in which $b= -1$). Every powers projection, therefore, when used along one axis only, always straightens out a curve whose formula involves that particular power of the one variable. We may treat all the possible

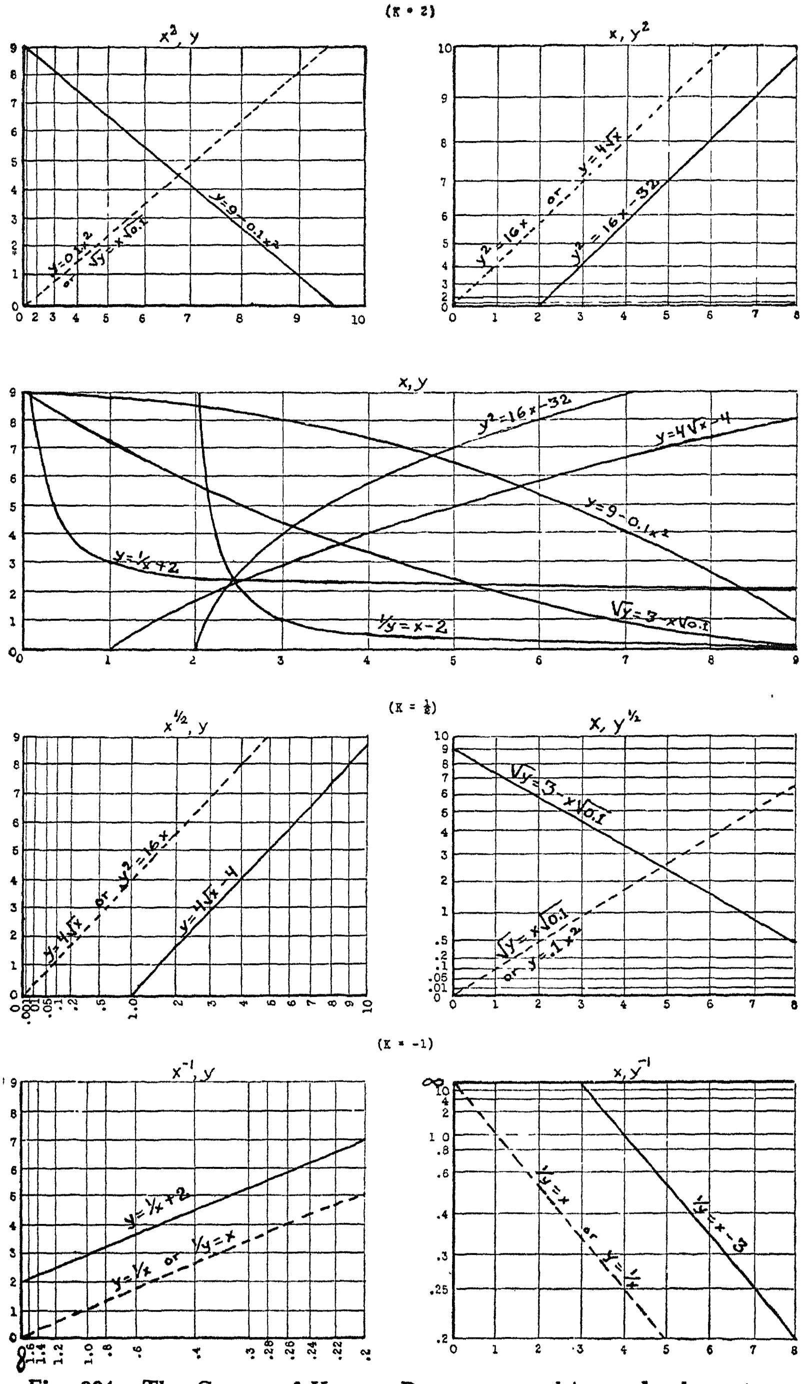


Fig. 394. The Curves of Known Powers, $y=ax^k+c$ and, $y^k=ax+c$. See footnote on opposite page.

powers projections as but one class, with formulas of the type $y = ax^k + c$. These formulas contain three constants, only two of which, however, are unknown. The third constant is known, and is the exponent of the variable. And surely it is clear that whenever the power of a variable is known, that power may be laid off upon the scale for that variable so that the plotting of only the first power of the variable (that is, the variable itself) thereon, will make the curve a simple linear one. The power remains in the scale and hence in the formula, but has vanished from the curve.

In all powers projections so far considered, we have used the special projection upon one scale only, the other being uniform. The equation being $y = ax^k + c$, it is clear that the x-scale has been specially projected, for the given power, k, of the variable, x. Care must be taken to keep this arrangement, for a reverse arrangement will fail to straighten out the curve. Only in the case when $c = 0$ and the equation reduced to $y = ax^k$, is it immaterial which scale be subjected to the powers projection, for here we may write $y = ax^k$ or $\sqrt[k]{y} = a'x$, in which $a' = \sqrt[k]{a}$. The line will therefore be straight either upon the powers projection of one scale or the corresponding root projection of the other. $y = ax^k$ is, however, too easily straightened out by the log projections, as we have seen, and hence the case is of no value. The real use for the powers (and roots) projection of the y-scale is in the wholly different equations of the form $y^k = ax + c$ (including $\sqrt[k]{y} = a'x + c'$).

Closely related to the projection of reciprocals, is the projection of products. Thus the equation $y = \frac{a}{x} + c$ may be written $xy = a + cx$. In this form we see that the equation will not only straighten out upon semi-reciprocal scales, $\frac{1}{x}$, y, but also

---

SCALE PROJECTIONS FOR FIG. 394:

Square of X.		*Square of Y.*
	Arithmetical	
Square root of X.		*Square root of Y.*
Reciprocal of X.		*Reciprocal of Y.*

Six typical curves are shown by full lines, one on each of the specially projected scales and all upon the arithmetical projection. They straighten out only upon the scales on which they are plotted. If, however, $c=0$, that is, there is no added constant and the equation is reduced to $y=ax^k$, then the curves are straight upon either of two different projections, thus on x^2, y or x, $\sqrt[2]{y}$; on x, y^2 or $\sqrt[2]{x}$, y; and on $1/x$, y or x, $1/y$. Such curves are shown by broken lines.

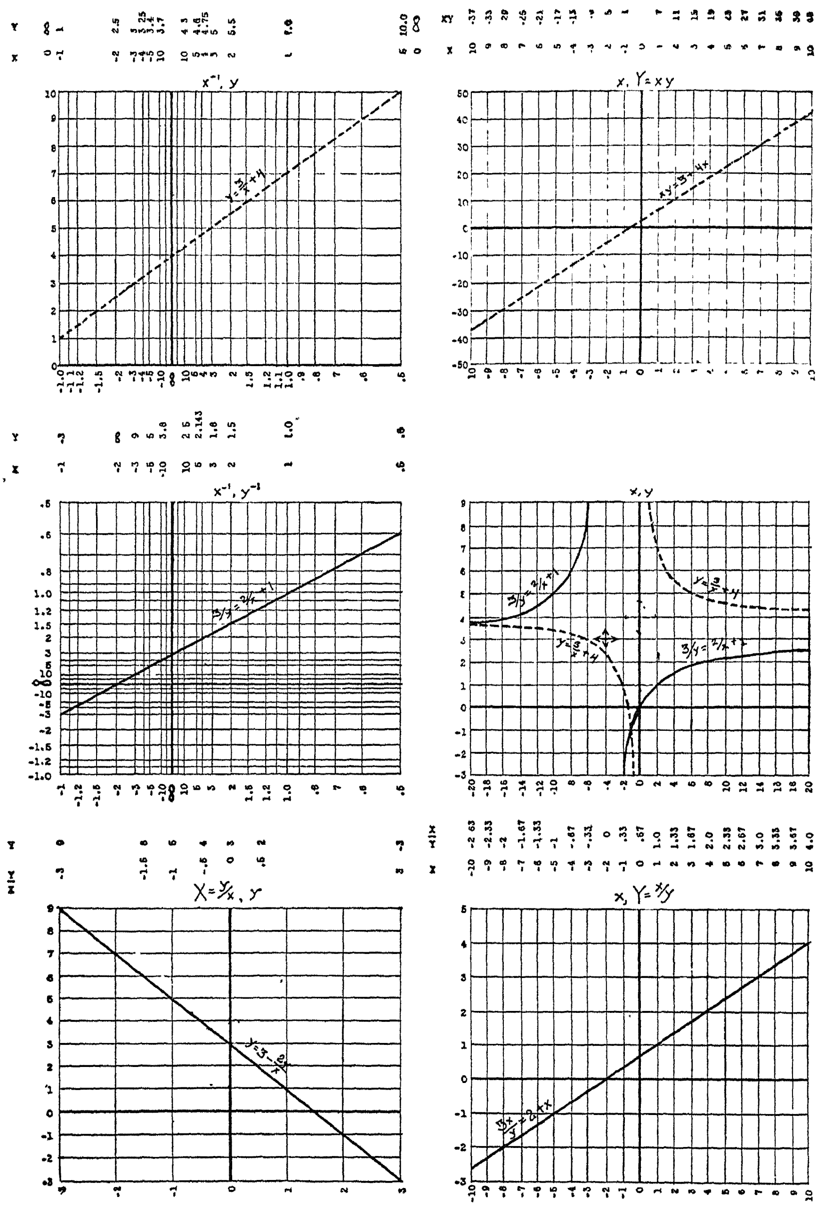


Fig. 395. **The Hyperbolic Curves,** $y=\frac{a}{x}+c$ **and** $y=\frac{x}{a+cx}$

See footnote on opposite page.

upon the semi-product scales, x, xy. Product and quotient scales are, however, a little hazardous, in that the introduction of one variable into both scales may often force a curve to approach nearer to, though not entirely to, a straight line, without the least real significance. Moreover, they require some computing, as the series of y original data must be replaced by the series xy, in which each value of x is multiplied into its corresponding value of y. To be sure, no special scale need be projected, the products or quotients being put upon a uniform (arithmetically projected) scale.

The most interesting use of the reciprocal projection is for the equation in which both variables are in reciprocal form, namely $y^{-1} = ax^{-1} + c$. This is the equation of the ordinary hyperbola. On uniform scales, it is asymptote to the co-

---

EXAMPLES AND SCALE PROJECTIONS FOR FIG. 395:

Single Reciprocal	*Product, xy*
$y = \frac{3}{x} + 4$	$xy = 3 + 4x$
Double Reciprocal	*Arithmetical*
$\frac{1}{y} = \frac{2}{3x} + \frac{1}{3}$	*Both*
Quotient, x/y	*Quotient, y/x*
$\frac{x}{y} = \frac{2}{3} + \frac{x}{3}$	$y = 3 - \frac{2y}{x}$

Two typical equations of hyperbolic curves are shown here. The one shown by the broken line, involves the reciprocal of one variable only and straightens upon a single reciprocal projection; from it a series of products of the two variables can be computed which straightens upon arithmetical paper on the plotting of x and xy. Note that its asymptotes are 0 and c; and that the value of c can be found by inspection at the intersection of the curve with the y-axis, calibrated as infinity, on the reciprocal projection. The value of a can be found from the phrase, $a = y - c$ when $x = 1$.

The other equation, shown by the full line, involves reciprocals of both variables and straightens out upon the double reciprocal projection, from it two quotient series can be computed which yield straight lines. Note that its asymptotes are $y = 1/c$ and $x = -a/c$ and can be read at the intersections of the curve with the axes, calibrated infinity, upon the reciprocal paper; from which the values of c and a can be easily found.

(Note.—The small letters over each diagram in this chapter show the functions (of the variables) plotted. Large letters, X and Y, indicate the scale-figures or calibrations, and are omitted if these are graduated for the same functions; the curve can then be plotted directly from the scales without finding the functions. The presence of the large letters indicates that the curve cannot be directly plotted from the scales, that the indicated functions must first be found and these (instead of the variables) must be plotted from the scale-figures.)

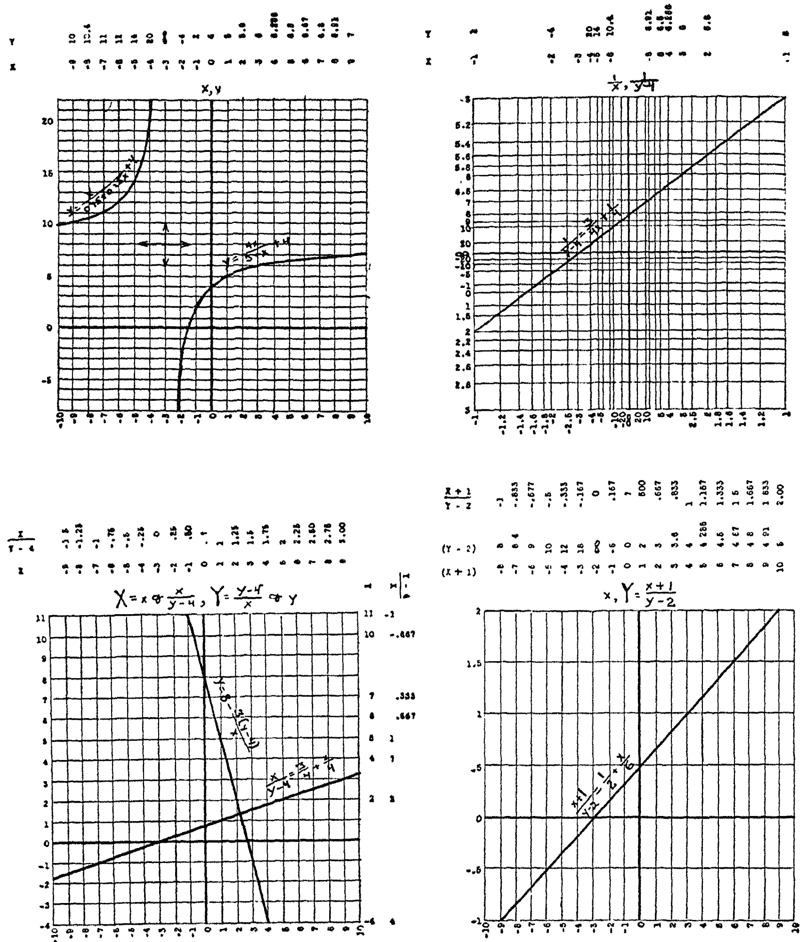


Fig. 396. The Hyperbola with Three Constants, $y = \dfrac{x}{a+cx} + d.$

SCALE PROJECTIONS:

Arithmetical. *Double reciprocal, one shifted.*
Quotients, singly shifted. *Quotient, doubly shifted.*

A single typical equation is shown in hyperbolic form on arithmetical scales and straightened upon reciprocal scales, one of which has a shifted zero. The reciprocal scale with the shifted zero can be prepared only when the added constant, d, is known, since the zero is shifted by its amount. Also when d is known the quotient series, $\dfrac{x}{y-d}$, or $\dfrac{y-d}{x}$, can be computed from the data and will yield straight lines upon the plot of either, x, $\dfrac{x}{y-d}$ or $\dfrac{y-d}{x}$, y. When d is un-

ordinates $x=\frac{a}{c}$ and $y=\frac{1}{c}$. From its form, $\frac{1}{y}=\frac{a}{x}+c$, we can see that it will straighten out upon paper in which both scales are reciprocally projected, $\frac{1}{x}, \frac{1}{y}$. If for any reason we desire a chart giving detail to different parts of the curve, we can re-write the equation as $\frac{x}{y}=a+cx$ and then we see that by the use of the scales, x, $\frac{x}{y}$, in which both are uniform and one, the y-scale, is used for the plotting of the quotients of the values of x by their corresponding y-values, the curve can again be straightened out. Or we can cast the same equation into the form $\frac{1}{y}-c=\frac{a}{x}$ or $1-cy=\frac{ay}{x}$ and so see that the curve will straighten out upon the scales, y, $\frac{y}{x}$. Here are three arrangements by which the ordinary hyperbola can be straightened out. Take your choice. The constants "a" and "c" can be found by inspection from the plotted curve on the double reciprocal paper.

If we write this equation in its usual form, $y=\frac{x}{a+cx}$, it will occur to the student that it may occur in modified form with a third added constant, thus $y=\frac{x}{a+cx}+d$. The curve of this modified equation will still be hyperbolic upon uniform paper, but will no longer pass through the origin of the chart. Just as in the first and third equations discussed, so here the added constant may be determined by the ordinate of the curve at the zero-point on the x-axis, that is, the point where

---

known, the zero must be shifted along both axes to some known point in the curve, and if we call the co-ordinates of this known point (it may be any we wish to select) x_0 and y_0, then we can compute the quotient series $\frac{x-x_0}{y-y_0}$ which yields a straight line y upon the plot of x, $\frac{x-x_0}{y-y_0}$. The value of d is obviously shown by the intersection of the curve with the origin, that is, $d=y$ when $x=0$. Note that the asymptotes, $y=1/c+d$, and $x=-a/c$, are shown upon the reciprocal projection by the intersections of the curve with the axes, calibrated infinity; from this the values of c and a are easily found.

the curve intersects the y-axis upon arithmetically projected scales. This cannot be read upon the reciprocal scales as the latter never reach zero, since zero would have to be plotted at its reciprocal, infinity, a manifest impossibility. We must therefore first plot the curve on uniform scales to determine, if we can, the value of d by inspection. Then if we write the equation in the form $y-d=\frac{x}{a+cx}$ and then $\frac{1}{y-d}=\frac{a}{x}+c$, we shall see that the curve will straighten out upon the reciprocal scales $\frac{1}{x}, \frac{1}{y-d}$. In short we have now come to the use of shifted or false zeros upon the reciprocal scale. The curve can also be straightened out by the quotient scales, $x, \frac{x}{y-d}$, and $y, \frac{y-d}{x}$. In every case there is a great deal of computing to be done.

When the value of d is not easily found, it may be advisable to use a method of differences, which has not so far been mentioned. Virtually this amounts to shifting the zero-point (or origin of measurements of the co-ordinates) to any convenient point we wish along the curve. To do this we first select a point upon the curve or an item in the co-ordinates which we may indicate by x_0 and y_0. Then we compute the difference between these co-ordinates and all other co-ordinates, x and y, in the data. Finally we divide the differences to get the series of quotients $\frac{x-x_0}{y-y_0}$ and plot these as co-ordinates over the abscissae of x. The reason for this is that the equation, $y=\frac{x}{a+cx}+d$, can be reduced, by subtracting the selected point, $y_0=\frac{x_0}{a+cx_0}+d$, to the form $\frac{x-x_0}{y-y_0}=a+cx_0+cx+\frac{c^2}{a}x_0x$. The added constant, d, which has caused all the trouble, has been, as you see, eliminated, and the constants, a, c, and x_0, are left. Write the second half of the equation as $(a+cx_0)+\frac{c}{a}(a+cx_0)\,x$, and you will see that the curve will straighten out on the plot of $x, \frac{x-x_0}{y-y_0}$. Of course this is not

a plot of the original series, it is only a plot of the cotangents[3] of the points upon the curve from the selected point in the curve, but if the derived curve be straight it is proof that the original curve has the formula $y = \frac{x}{a+cx} + d$.

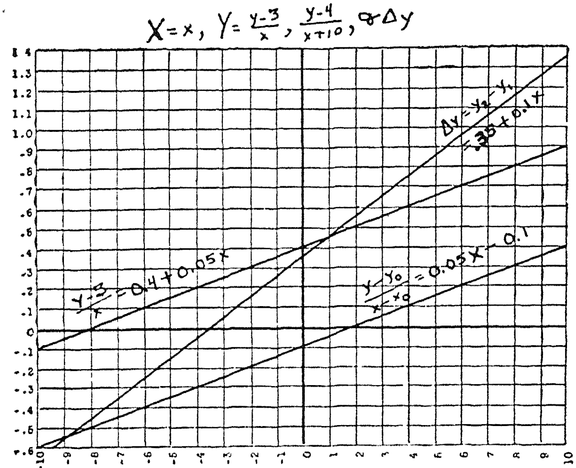
In all reciprocal scales the constants a and c are easily found by inspection. The axes of the scales are always calibrated as infinity along each scale and cross the curve at its asymptotes. The asymptotes have the value $x = \frac{a}{c}$ and $(y - d) = \frac{1}{c}$. So we need only substitute the observed values for x and y to obtain the constants and complete the formula.

Similar methods can be used to straighten out the ordinary parabola, the equation for which is $y = a + bx + cx^2$. (The reader will note that to arrange the variables in ascending order, we have altered the symbols for the constants, and that this equation is really a modification of the simple linear one, $y = ax + c$.) A little study will show that the value of the constant a is the value of the curve at its intersection, experimentally drawn, with the y-axis. Since the equation can be written $\frac{y-a}{x} = b + cx$, it is obvious that if we know a, we can compute the series of quotients $\frac{y-a}{x}$ and that the curve of these will straighten out upon the scales, x, $\frac{y-a}{x}$. If we do not know the value of a, we can use the method of shifting the origin to a point upon the actual curve[4] and will get a straight line by plotting x, $\frac{y-y_0}{x-x_0}$ (not $\frac{x-x_0}{y-y_0}$ because reciprocals are not involved in the equation). If the values of x in our data form an arithmetical series, we can take the successive differences of the y-values, that is, $\triangle y$, and will find that the plot of x, $\triangle y$, is a straight line.[5] The polynomial involving higher powers sych as $y = a + bx + cx^2 + dx^3 + \ldots$ must be successively differentiated or the method of determinants

[3] Cotangents only because they are reciprocals.

[4] By substituting $y_0 = a + bx_0 + cx_0^2$ we get $y - y_0 = b(x - x_0) + c(x^2 - x^2_0)$ or $\frac{y-y_0}{x-x_0} = b + cx_0 + cx$, in which the phrase $(b + cx_0)$ is a constant.

[5] The formula for $\triangle y$ is to be found in Lipka, "*Graphical and Mechanical Computation*," p. 146.



		-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10
ΔY			-.55	-.45	-.35	-.25	-.15	05	05	15	25	.35	45	55	65	75	.85	95	1.05	1.15	1.25	1.35
$\frac{Y - Y_0}{X - X_0}$	$\frac{Y - 4}{X + 10}$		-.55	-.50	-.45	-.40	-.35	-.30	25	20	15	-.10	-.05	0	.05	10	15	20	25	30	35	.40
$X - X_0$	$X + 10$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$Y - Y_0$	$Y - 4$	0	-.55	-1.0	-1.35	-1.60	-1.75	-1.80	-1.75	-1.60	-1.35	-1.00	-.55	0	.65	1.40	2.25	3.20	4.25	5.40	6.65	8.00
$\frac{Y - A}{X}$	$\frac{Y - 3}{X}$	-.10	-.05	0	.05	.10	.15	.20	25	.30	.35		45	60	55	60	65	.70	75	80	85	90
$Y - A$	$Y - 3$	1	.45	0	-.35	-.60	-.75	-.8	-.75	-.60	-.35	0	.45	1.00	1.65	2.40	3.25	4.20	5.25	6.40	7.65	9.00
Y		4	3.45	3	2.65	2.40	2.25	2.20	2.25	2.40	2.65	3.00	3.45	4.00	4.65	5.40	6.25	7.20	8.25	9.40	10.65	12.00

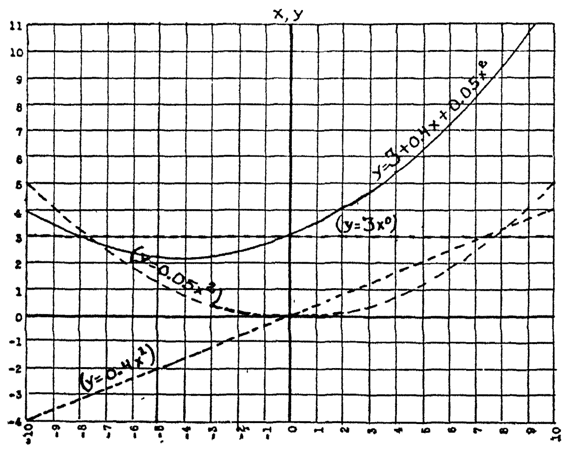


Fig. 397. The Parabola, $y=a+bx+cx^2$.
See footnote on opposite page.

used, but the mathematics in such work is outside the scope of this volume. Indeed, the phrase cx^d may be added to any of the foregoing equations and will produce the same difficult results.

The close observer may note that so far no mention of the semi-logarithmic chart has been made. The curve which straightens out upon it is known as the simple exponential or logarithmic curve. The formula of such a curve is of a wholly different nature from any so far considered, for in the latter all exponents have been constants. In the equations of exponential curves, we meet with variable exponents. The simple exponential curve has the formula, $y=ab^x$. The curve upon arithmetically projected scales is convex to the x-axis, rising rapidly as x increases positively, and having a horizontal asymptote to the x-axis as x increases negatively and intersecting the y-axis, at the value of a, since "b^x" becomes unity at the y-axis (x being zero there and the zero power of any number being one). Now if $y=ab^x$, then of course $\log y = \log a + x \log b = a' + b'x$ (in which a' and b' are constants, the logarithms of a and b). The second part of the equation becomes familiar enough in this last form, being the very first type considered. So if we plot the curve upon semi-logarithmic paper, x, $\log y$, it will straighten out. And conversely all straight lines upon semi-log paper have the equation $y=ab^x$. And this formula applies to all historical data which have straight-line curves upon semi-log paper. It is the formula of the law of organic growth.

Several variations of the exponential curve equation are obviously possible. If a constant be added we have $y=ab^x+c$. This we turn into $y-c=ab^x$ and from it we derive $\log (y-c) = \log a + x \log b$ or $\log (y-c) = a' + b'x$. Here we see a need for the shifted zero upon the logarithmically projected scale, an the curve straightens out upon the scales x, $\log (y-c)$. If c is unknown it can be computed from the experimental curve

---

SCALE PROJECTIONS FOR FIG. 397:
Quotients (shifted) and differences.
Arithmetical.

The ordinary parabola (shown by a full line on the lower diagram) is the sum of three different curves (shown by broken lines) and cannot be straightened out upon any useful projection. It can be made to yield, however, various straight lines for series which have been computed from its known data and these afford a test of its equation.

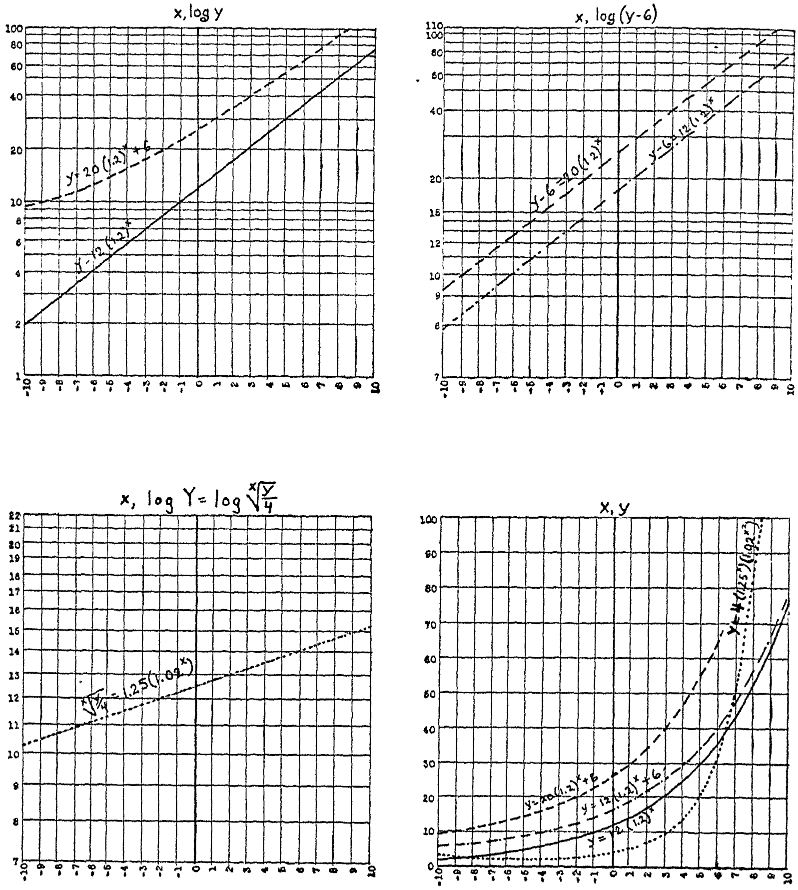


Fig. 398. Exponential or Logarithmic Curves.

EXAMPLES:

$y=ab^x$, or $\log y=\log a+x \log b$

$y=ab^x+c$, or $\log (y-c)=\log a+x \log b$

$y=ab^xc^{x^2}$, or $(1/x) \log (y/a)=\log a+x \log b$

The simple exponential equation $y=ab^x$ straightens out upon semi-log paper. When a constant c is added, the zero must be shifted by the amount of the constant on the log axis; one example of this (shown by the broken line), is plotted upon three diagrams to show its behavior, another which on arithmetical paper parallels the simple equation shown, is plotted on two (shown by the dot-and-dash line). These curves are hyperbolic. When a higher power of the variable is added, the curve becomes parabolic and cannot be straightened. A quotient series with shifted zeros can, however, be computed, namely, $\dfrac{\log y-\log a}{x}$, which will yield a straight line; this line (shown dotted), is here plotted upon semi-log paper as $\dfrac{\sqrt[x]{y}}{a}$ simply for the sake of variety.

upon uniform paper.[6] Another exponential curve is that which has the formula $y = ab^x c^{x^2}$. If it be written as $\log y = a' + b'x + c'x^2$ (in which the primes of the constants again represent their logs) we see that it is very similar to the ordinary parabola, and indeed, precisely the same methods must be used to straighten it. If a is known, we can plot it upon the uniform scales x, $(\frac{\log y - \log a}{x})$. If a is unknown we can use either of the other two methods, and derive series whose curves straighten out upon the scales x, $(\frac{\log y - \log y^o}{x - x_0})$ and x, $\triangle \log y$. Other exponential curves have still other formulae, which are often but modifications of any of the foregoing through addition of other variable powers, such as d^x in the equation $y = a + bx + cd^x$. These more complicated equations must be subjected to even more devious calculations before derived series can be found which straighten out and prove the equation.

The reader should not consider from this brief summary of the scale projections which straighten out non-periodic curves, that all or even nearly all curves can be straightened out by them. And the non-mathematical reader will doubtless have a wholesome respect for the processes of curve equating even by the above methods. He will probably find little difficulty with the simple linear, the simple parabolic and hyperbolic, and the simple exponential curves, requiring as these do only the arithmetical, logarithmic and semi-log charts. But some curves are immensely difficult to express in equation form, and must often be broken into parts with separate equations for each part. It is true that these parts can be collected with proper mathematical symbols of limits, into a single equation and in this sense it is true that an equation can be written to any curve in the world.

But the long and complicated equation has little value. The equation for very irregular curves—such as the profile of a man's face—may take up more space than the curve itself. The disadvantages of complicated formulae are many. For one thing, a very complicated formula is difficult to understand

[6] The same method for finding c can be used as before, for the equation $y = ax^b\ c$.

even when it has been stated—the average person still has to plot its curve to understand its meaning. For another thing, very complicated formulae suffer from the danger of being made unnecessarily detailed or intricate by chance variations in the observations which form the data.

This last consideration, the danger of chance variations in the observed data, leads us to the thought that the "true curve" for the data, if all errors were absent, might be a very simple curve, easily expressed by an equation, while the curve of the actually observed data remains irregular and complicated. We therefore oftentimes have to be satisfied by simple curves which closely approximate the actual curves, when such simple curves can be found. And the problem then becomes one of "fitting curves" with the best possible (that is, the closest fitting) straight lines, in the attempt to find simple and approximate descriptions and equations.

The reader will have seen by this time that much of the care expended on proper curve plotting has for its purpose the clear visualizing of the phenomena, but that still other care is expended in the attempt to capture the curve in a symmetrical or regular formation. And he will now see that one of the chief purposes of symmetry and regularity is to enable us to formulate laws governing the behavior of the phenomena represented by our data and curve. In the discussion of fitted straight lines, which is so far as it seems desirable to enter the subject in this book, he will be reminded of the "trend" and "secular change" discussed previously in historical curves; in fact for historical series the secular trend is often considered to be a fitted straight line. And he will now also see that these secular trends can be expressed mathematically in equations. He will also see that the operations of interpolation and extrapolation can be even more precisely performed when the equations are used than with charts only. He will see, in short, that the possibilities of mathematical description or summarization of curves opens up to him a valuable adjunct to the use of the curves themselves.

PART V. CALCULATING CHARTS

CHAPTER XLV

CURVES FOR FORMULAE

Having seen something of the way in which formulae or equations can be written to curves, we can reverse the process and prepare curves to illustrate formulae. In this way, we no longer seek the mathematical statements describing a curve, but we seek the curves illustrating a mathematical statement. The advantage of writing an equation to a curve lay in the fact that, from the equation alone, we could, by mathematical operations, find the values represented by each or all of the plotted points along the curve; the advantage of drawing the curve to illustrate an equation lies in the fact that without bothering about the mathematical processes, we can read the values represented by the equation directly at a glance from the chart. In short, the chart may be made a substitute for the processes of calculation and computation, and the chart then becomes a calculating machine.

If, as in the previous chapter, we have a curve for which the mathematical equation is $Y = 2X + 3$, and we wish to find the value of Y, when X, let us say, is 5, we do not have to solve the equation by mathematical processes, multiplying 5 by 2 and adding 3, but from a glance at the chart we can see the Y-value of that point on the curve whose X-value is 5. We follow the ordinate from 5 on the x-scale up to the curve and from the intersect point (where the curve passes through or intersects the ordinate) we follow the abscissa or horizontal to the y-scale and read 13, the answer. In this case, it is true that the mathematical operation of solving the equation seems simpler than the graphic one for the reason that we have selected for illustration of the principle a simple mathematical equation. But you will find many complicated formulae and equations in which the mathematical operations are far more tedious and lengthy than the graphic process. In such cases it will be useful for you to be able to construct calculating

curves and charts by which mathematical equations of the given type can be readily solved.

The purchasing agent, perhaps, buys in foreign markets and must multiply his quotations by the prevailing rate of foreign exchange and add perhaps certain local charges in this country, before he can compare the values of different offers. To interrupt telephone conversations with these mathematical operations would perhaps be difficult, but he could be provided with a special chart on which he would see at a glance the real value of offers without interrupting his telephone conversation to the parties concerned.

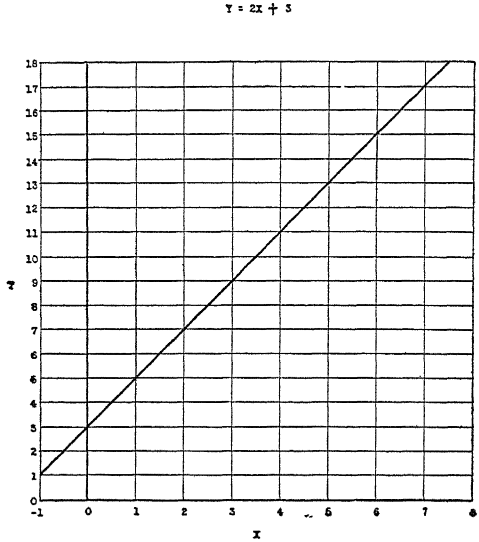


Fig. 399.

That a single straight line curve upon an arithmetically projected chart-field will illustrate a simple mathematical equation involving only two variables in the first degree, we already know, for any straight line upon arithmetically projected chart-fields has an equation of the general form $Y = aX + c$ (in the right side of which a and c are given constants and X alone is variable). We can, however, by a series of such straight lines show the equation for two independent variables. Let us assume for example that c is a variable and call it Z and that the constant a is 2. In other words let us prepare a calculating chart for the equation $Y = 2\,X + Z$. As we have seen in the last chapter the figure 2 determines the

slope of the straight line curve and if the x-scale is only half as great as the y-scale then the slope of the straight line would be rigidly 45° to the x-axis of the chart. The added element

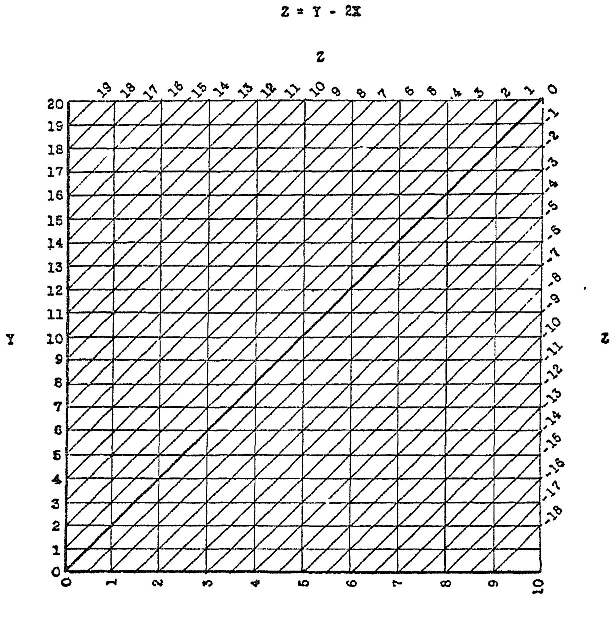
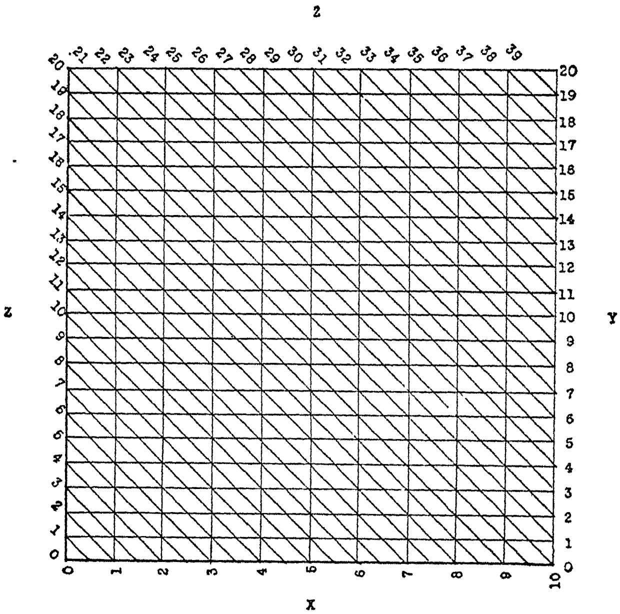


Fig. 400.

Z merely determines the height or position of the straight line curve upon the chart; the straight line curve passes through the origin of the chart when Z is 0 and in general intersects the y-axis at the value of Z because at the y-axis the value of X is 0 and the equation is $Y=Z$. Now because Z itself is a variable we cannot show the equation by a single straight line but must use a series of straight lines, each for different values of Z and must therefore mark off a scale of Z upon the straight line curves themselves. The result is a chart with a series of parallel straight line curves which are diagonal upon the chart and enable us at once to find the values of Y when $Y=2X+Z$. To read a certain value, as, for example, when X is 5 and Z is 3, we need merely read up the ordinate from the point 5, on the x-scale, to the diagonal line or curve marked 3 on the z-scale, and from the intersect of this particular curve with the ordinate, read horizontally across the abscissa to the point on the y-axis where we find 13, the answer.

To use this chart for subtraction is very easy, for we merely reverse the process and the dependence of the variables, saying that if $Y = 2X + Z$, then $Z = Y - 2X$. If $Y = 10$ and $X = 2$, then we read across the abscissa from the point of 10 on the y-scale to the ordinate from the point 2 on the x-scale, and note the value of the diagonal which passes through this point, namely 6 on the z-scale. It is of course not necessary to use whole numbers either upon the chart or in the equation for we can easily interpolate between the

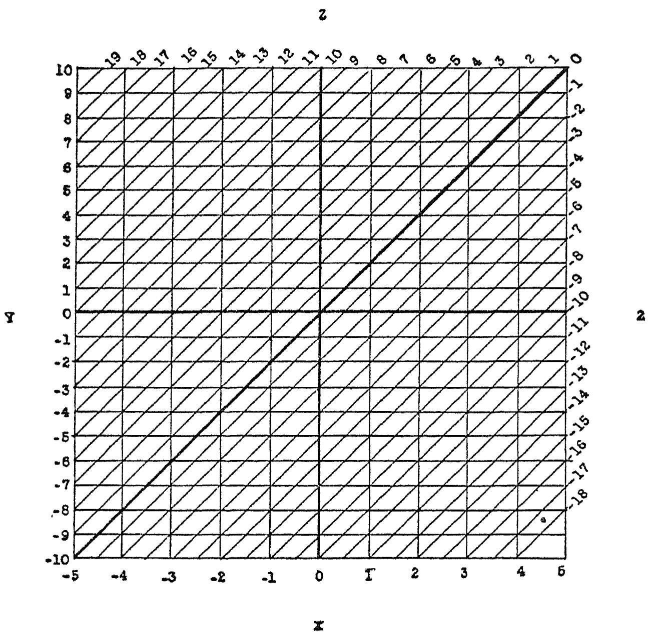


Z = 2X + Y

Fig. 401.

actual ruling on the chart to estimate, very closely, the values desired, when they are fractional. These subtractive charts can also be made to show the difference, not upon the straight line diagonal curves of Z, but upon the y-axis itself, by making the curves express the general equation $Y = Z - aX$ and making the diagonal curve a descending instead of an ascending one, as illustrated in the previous chapter. Still another method for obtaining the same result would be to carry the chart described in the last paragraph down into the negative side of the x-axis.

Indeed the calculating chart only becomes difficult to understand when we begin to talk about it. The simple chart is much more easily made than described. Yet it is necessary,



Z = Y - 2X

Fig. 402.

for charts of the more complicated formulae, that the elements which go to make up the simple chart be clearly defined. And the first consideration of importance is the distinction between physical distances upon the chart and numerical values assigned thereto. If we call the two axes of the chart (x), and (y) and the diagonal dimension (z), then we have at least distinguished three different possible places in which scales may be projected physically and given numerical values. If we indicate the physical distances along these scales, measured from an origin-point, as x, y, and z, respectively, and the numerical values finally assigned to these distances (i.e., the scale figures) as X, Y, and Z, we have a simple means of keeping two more details separate in our minds.

The importance of distinguishing in this way between final calibrated values or scale-figures, X, Y, and Z, and the

actual plotted scale-distances x, y, and z, cannot be underestimated, for confusion at this point will baffle the student for the remainder of his work upon calculating charts. It is to be understood that the small letters, x, y, and z, are merely essentials in the planning and making of charts; they do not appear upon the finished work. It is to be understood that the large letters, X, Y and Z, are merely symbols for the different variables in the equation to be calculated. If these variables are indicated by other symbols in the equation, then the large letters will not appear on the finished work, but the accustomed symbols will be substituted for them. The large letters are useful in planning the work as they clearly indicated the axis or scale upon which the variables will appear. If no better symbols are to be had, then the large letters X, Y, and Z, one or all, may be retained upon the final chart and its formula. Indeed, even the small letters, x, y, or z, may be finally used for this purpose; they then of course indicate the variables and scale-figures and have no application to scale-distances. But during the stage of making the chart, we shall always use the symbols consistently in the meanings specified. Thus if we have the equation "income − operating expense = operating profit" or "$i - e = p$," we shall substitute Y, let us say, for i; Z for e; and X for p; and write $Y - Z = X$. When the chart is finished, we shall substitute the original symbols again, and write $i - e = p$.

But between the scale distances, x, y, and z, and the final calibrations, X, Y, and Z, an elaborate structure of modifications and substitutions may be built up. These are necessary for very complicated formulae; in simple equations they fall together like a house of cards and can be wholly disregarded. Thus if our equation be $Y = X + Z$, we can obviously lay off the distances, x, y, and z, directly from the equation. We may even use the same face of the ruler for both y and x, that is, along the axes of the chart, plotting the diagonals to conform to the equation. But when $Y = 2X + Z$, the chart becomes very tall, and as we have seen, it is just as well to lay off the Y and X scales differently.

Since we have occasionally in this way to use different units of measurement in laying off scales, it is well to have clearly in mind one common unit of measurement for the entire chart. This unit we call the "modulus" of the chart; and it does not matter whether the modulus be one inch, one foot,

one centimeter, or any fraction of these, so long as it be the same for all parts of the chart the proportions of the various parts of the chart are the same. The modulus then is simply a general unit of distance which serves in planning the chart to equate the scale distances and the scale values; thus, $x=mX$, or $x=2mX$.

Now it is a great convenience to plot distances directly from the data, that is, the values or scale-figures to be assigned. When this can be done we can copy scale-figures directly from our ruler as we plot. And here secondary moduli for each scale become useful. These are merely fractions or multiples of the chart-modulus, and when they differ from the latter, may be indicated by m_y, m_x, or m_z. Thus when $x=mX$, $m_x=m$, but when $x=2mX$, $m_x=2m$. In the charts already considered, we have seen the chart of $Y=2X+Z$ made with the horizontal units of measurement twice as long as the vertical ones. If the vertical units be m, then the horizontal ones are $2m$.

Experience will show that it is best to proportion a chart in such a way that all intersections be as sharply drawn as possible. The object is to make readings from the chart accurate. If two lines are perpendicular, there can be little doubt about their intersection point, but when they cross at small angles (that is, are nearly parallel) it is not so easy to decide the exact point of intersection. Since the x and y co-ordinates are perpendicular, obviously the z-diagonals cannot cut both co-ordinates more sharply than at 45 degrees. So the most desirable form of chart is one in which the z-diagonals form about 45° angles with the axes. And it is the primary purpose of the scale-moduli (not the chart-modulus) to produce this condition. When $Y=X+Z$ and x and y have equal moduli, the z-diagonals, as we know, have the right slope. And so when $Y=2X+Z$ it is easy to see that the modulus of the x-scale (letting $m_y=m$) must be $m_x=2m$, before the diagonals will have the same slope. Here we may note that $\frac{m_x}{m}=2$, the coefficient of X in the equation. And it is a useful empirical rule that the coefficients of the variables (on the axes, that is, X or Y) are the ratios of their scale moduli to the chart modulus.

The scale moduli (as distinct from the chart modulus) serve still another purpose, for since they form what we might

call "plotting instructions," they can be used to indicate the side of the engineer's hexagonal rule[1] which is to be used. Thus if we use a chart modulus of one inch, we can plot m from the 10-side of the rule, $\frac{1}{2}m$, from the 20-side of the rule, $\frac{1}{3}m$ from the 30-side, and so on.

From the outset in chart making for formulae, we must keep in mind the desirable limits of the variables to be shown by the scale figures. If our chart is to be used in calculating a few pounds, it would be foolish to make it include tons as well, for then the scale for pounds would be so small that it could not be accurately read. If the price of paper is quoted in cents, why make a chart which shows millions of dollars, and on the scale of which cents are so small as to be invisible. Obviously the larger our scale becomes the more clearly it can be read, and the more accurate will be its calculations. Hence we should try to include in the range of the scale only its useful parts that we may make them as large as possible. This calls for the setting of limits f r the range, an entirely arbitrary matter, for which it is only necessary that we know the extreme high and low values of the variables which will be met with in the use to which the chart will be put. Having determined these values of the independent variables, we can write them into our formula by a convenient trick, thus $y=x\left|\begin{matrix}10\\0\end{matrix}\right|+z\left|\begin{matrix}12\\5\end{matrix}\right|$ and hence, in full $y\left|\begin{matrix}22\\5\end{matrix}\right|=x\left|\begin{matrix}10\\0\end{matrix}\right|+z\left|\begin{matrix}12\\5\end{matrix}\right|$. Now we know how much space to give to the chart, or how large to make the chart-modulus for a chart of a given total size.

We are now in a position to consider the havoc wrought by constants in a given formula for which we are making a chart. If these constants be coefficients of the variables, we have an equation of the type $bY=aX+cZ$, we shall have the scale moduli, $m_y=bm$, and $m_x=am$. The scale modulus of the z-scale for diagonals need not be calculated, Z is much more easily entered upon the chart from observations of the actual values for various points after the co-ordinates have been calibrated. The formula $bY=aX+cZ$ can be written $Y=\frac{a}{b}X+\frac{c}{b}Z$; this will enable us to make $m_y=m$ and $m_x=\frac{a}{b}m$, which may

---

[1] For a description of the engineers' rules or scales, see Chapter XVII.

give an easier plotting scale directly from the ruler. Thus if we have $14Y = 7X + 3Z$, it is a convenience to plot $y = mY$ and $x = \frac{7}{14}mX = \frac{1}{2}mX$, for we can plot and calibrate directly from the 10 and 20 sides of the ruler; but if we have $5Y = 2X + 3Z$, or $\frac{y}{2} = \frac{x}{5} + .3z$ it is more convenient to plot $y = \frac{1}{2}mY$ and $x = \frac{1}{5}mX$, for we then use the 20 and 50 sides directly.

Of course these considerations are largely directed at the simple co-efficients, but they hold also for more complicated ones. When we have an equation such as $157 = 37.295X + Z$, no rulers will serve directly and it would not pay us to plot $x = \frac{37.295}{157}mX = .237\ mX$ from a specially constructed scale (best secured by the method of triangulation[2]), instead we need only plot $x = .25mX$, which we can do from the 40-side of the rule, and shift the direction of the z-diagonals a little. When constants are added in the equation, the effect is not to enlarge or diminish the size of scales, nor to alter the scale-moduli in the least, but it is to shift the scale numbers, without otherwise disturbing them, along the axis. No matter how many constants be added, they can of course be lumped into one, thus $bY = aX + cZ + k$. Obviously the correction for the constant must be made upon one or another scales, that is the constant must be attached to one or another variable, to get it into the chart. Thus $(bY - k) = aX + cZ$, $bY = (aX + k) + cZ$, and $bY = aX + (cZ + k)$ are all forms of the same equation. The amount of shift is proportional to the amount of k, but care must be taken to divide it by the coefficient, if there be any, of the variable to which it is attached. Since $aX + k = a(X + \frac{k}{a})$, we must make $x = ma(X + \frac{k}{a})$ and if we wish to shift the scale (i.e. add the constant) after the scale (i.e. $x = amX$) has been plotted we must shift it by the amount of $\frac{k}{a}$, not of k. It is sometimes simpler to make the correction while plotting, that is plot for $x = m(aX + k)$ or $x = am(X + \frac{k}{a})$, directly by sliding

[2] For a description of the amplifying or diminishing of scales by triangulation, see Chapter XVII.

the ruler along until the calibration X is at the point of $(X+\frac{k}{a})$.

We have so far considered only one form of calculating chart the distinct feature of which is the parallel straight line z-diagonals. This is the chart for all equations involving the sum or difference of two variables of the first degree. It may be called, therefore, the additive chart. It is by far the most important. and useful, as well as the simplest chart of its kind. Moreover, it is the basis for so many other calculating charts that we have dealt with it in great detail, almost all of which will be essential to an understanding of the other types of calculating charts.

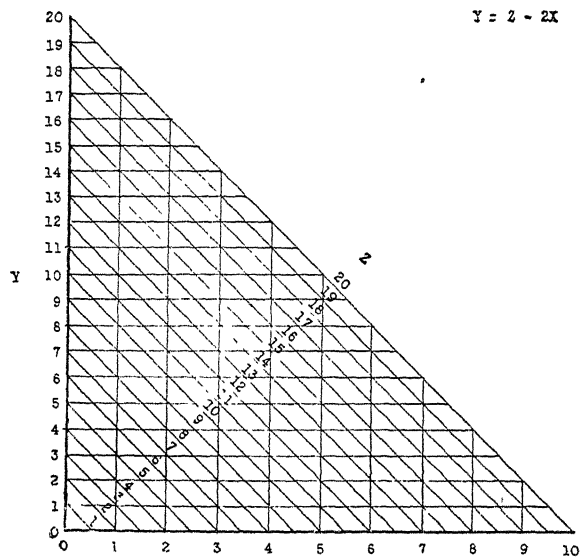
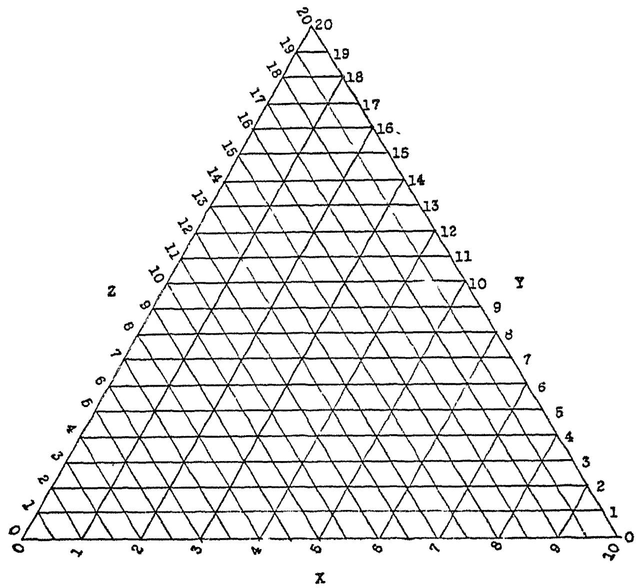


Fig. 403.

The reader may note, however, that the use of the scale-modulus for the projection of the z-scale has been expressly enjoined. This is a peculiarity of the rectangular chart, the z-scale being best laid off by inspection in it. The reason for this is that the scale of the z-diagonals is not easily commensurable with the x and y-scales. The diagonals form angles of 45 degrees with the other co-ordinates, when the scale-moduli of the x and y-scales are precisely adjusted; they form approximately the same angles when the adjustment is not complete

but is fairly close. Now if we could lay off all three sets at precisely equal angles the z-scale would become easily commensurable, and the scale-modulus of the z-scale could be used like the other scale-moduli.

There is much to recommend such an arrangement of trilinear co-ordinates. All intersections would be distinct, hence greater accuracy would be achieved in the use of the chart. The useful portions of the chart would be more compactly



Y = Z - 2X

Fig. 404.

positioned, hence space would be conserved and greater detail available. The form would be unusual and more attractive, an important feature, since, as we shall presently see, these charts are more pictorial and popular than business-like. Yet in spite of these advantages, the equilateral and equi-angular form is seldom or never used, probably because the average chart-maker has become so accustomed to rectangular co-ordinates.

The additive chart which we have described can be turned into a factorial one by logarithmic projection of scales. It can then be used to calculate the formula $Y = kX^aZ^c$, since log

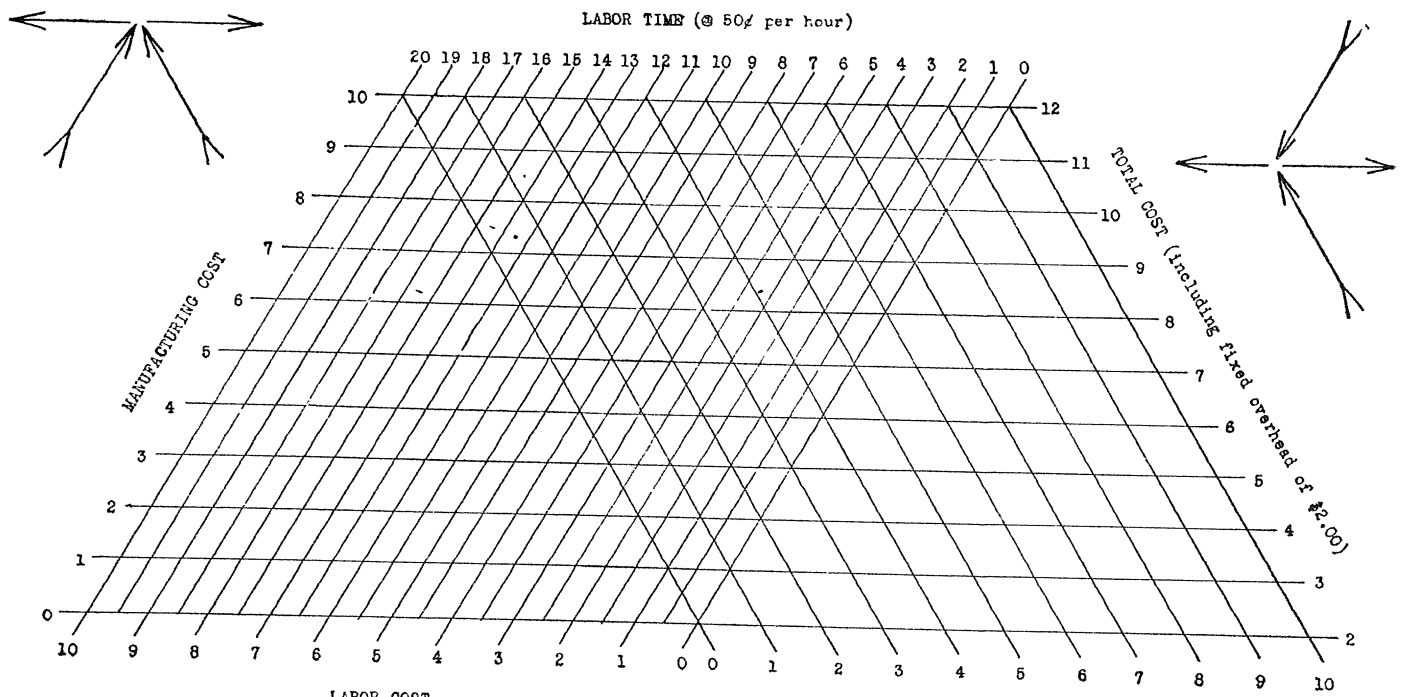
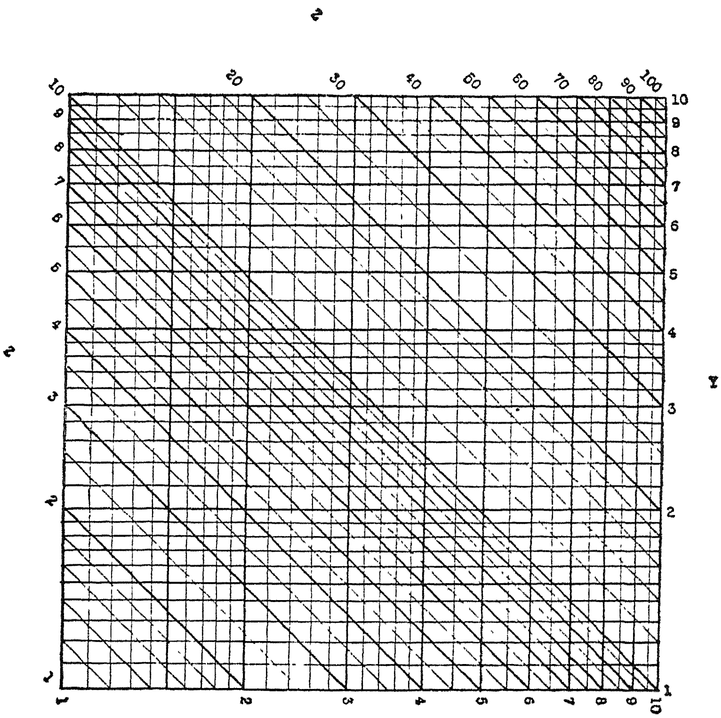


Fig. 405. A Simple Calculating Chart.

Used for finding the total cost of goods in a concern in which the labor cost is 50 cents per hour, with a total fixed overhead of $2.00 for each article. The arrows show the procedure in finding solutions.

$Y = \log k + a \log X + c \log Z$. The chart is prepared precisely as is the additive chart, and is much more generally useful.



X

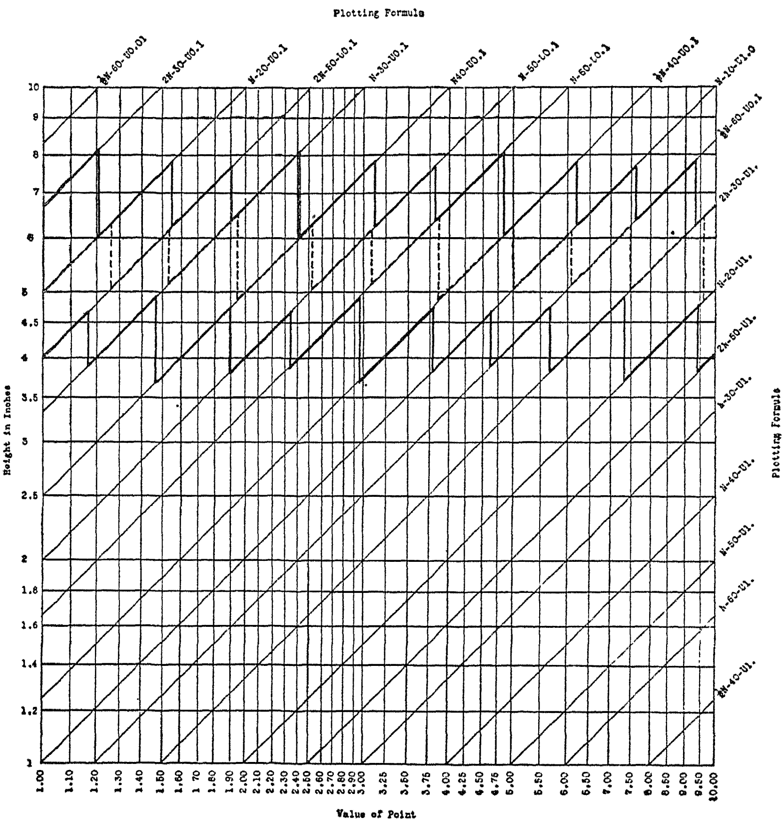
Z = XY

Fig. 406.

An exponential chart can be obtained by the combination of logarithmic and arithmetical scales, for the formula $Y = a^X Z$ and the like. Other possible combinations will occur to the student. The chief use of the chart, however, has so far been in its additive and factorial forms, having arithmetical and logarithmic scale projections only. With these two the student should be thoroughly familiar.

We come now to another chart which has straight-line *z*-curves or diagonals. It is easily distinguishable from the foregoing from the fact that the *z*-curves are not parallel to each other, but radiate from a common intersection point. The parallel line chart was simply a multitude of curves for the linear equation, $Y = aX + C$, in which many values of C were taken and C itself treated as a variable. The radiating straight-

line chart is simply a multitude of curves for the same linear equation, save that many values of A are taken, and A is

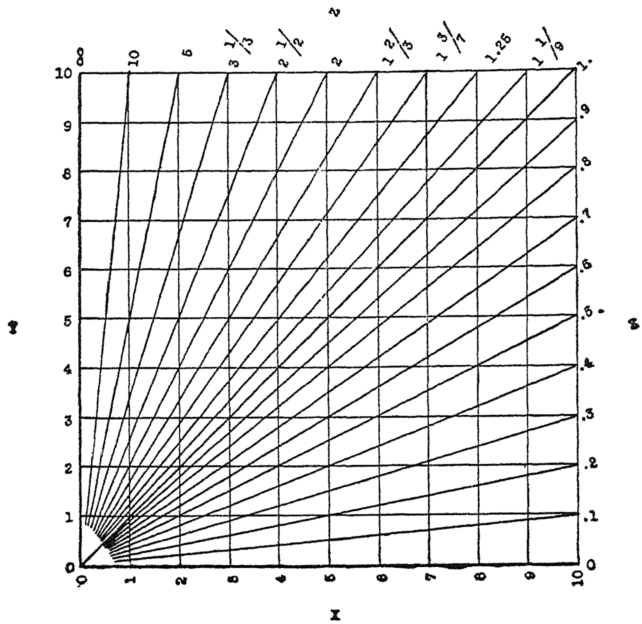


KEY TO SCALES FOR CHARTS

Fig. 407. Chart for Determining Scales of Curve-charts.
At the bottom of the chart find the value of the highest point in the curve. At the left-hand side find the height which you wish to give it on the chart. At the nearest intersection of the ordinate with a diagonal at this height read (in the scale of diagonals above and to the right) the side of the rule, N, and the plotting value of each unit on this scale (U).

treated as a variable. For convenience we will still call this third variable Z, and the equation therefore which this chart solves is of the form $Y=ZX$, when the common point through which all diagonals pass is the origin of the x- and y-axes. When the common point is located elsewhere on the y-axis, as at the co-ordinates $x=0$, $y=c$, the chart solves the equation $Y=2X+c$. The chart is primarily factorial, though it has

arithmetically projected x- and y-scales. The z-scale is a scale of angles, a circular function of the x- and y-scales.

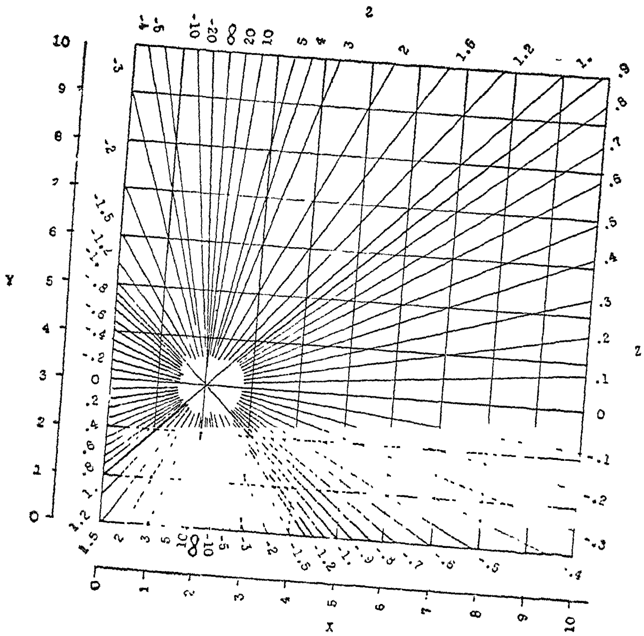


$$z = \frac{y}{x}$$

Fig. 408.

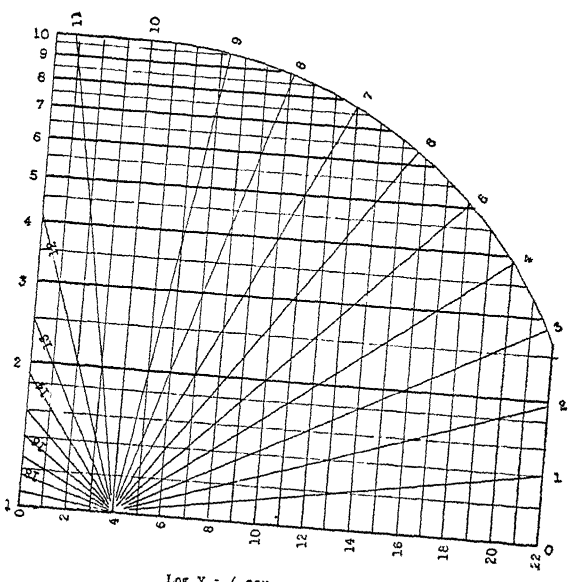
We shall not go into the principles of this chart in detail, it can be easily made, the x- and y-scales being arithmetical and the z-diagonals and z-scale being best put in by inspection. The chart is not of much value (save for one particular purpose), it is difficult to use accurately when the values must be interpolated between co-ordinates and diagonals and loses detail as the diagonals converge. For the processes of multiplication and division which are the main purpose of this arithmetical factorial chart, the logarithmic factorial chart is in every way, save one, the more satisfactory.

There are a great many other calculating charts upon co-ordinates, which have not the straight line diagonals or z-curves, but in which the third variable is shown literally by a series of curved lines, each bearing a particular value of Z. These are, of course, only multiple parabolic, hyperbolic, or other



$Z = \frac{Y - 3}{X - 2}$

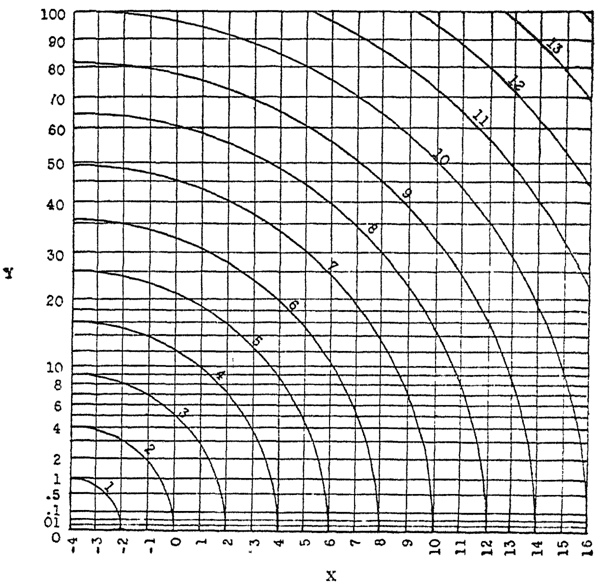
Fig. 409.



Log Y = (.05X - .2) tan 9Z

Fig. 410.

curves (even including circles and ellipses), and solve more complicated equations than the forms already discussed.

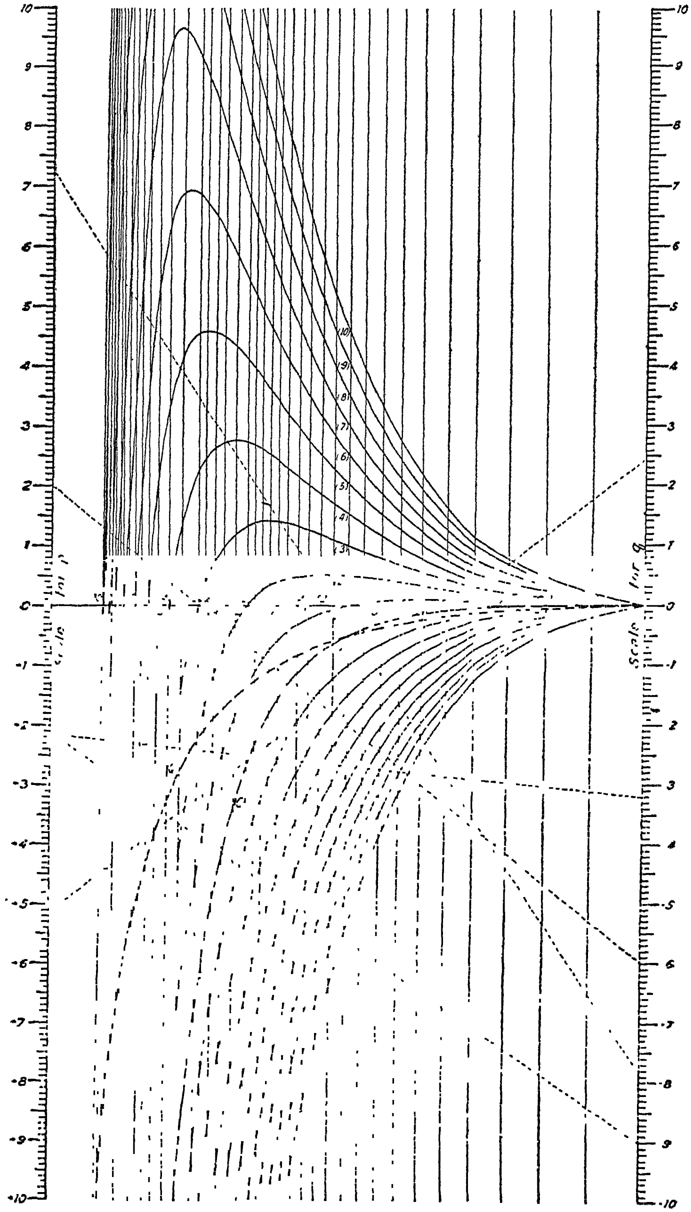


$$Z = \sqrt{\left(\frac{x}{2} + 2\right)^2 + y}$$

Fig. 411.

They are largely of academic interest, however, forming interesting exercises for the student and somewhat highly pictorial displays of the behavior of phenomena to which their formulae apply. For practical purposes they are of little value, since they take up much time and effort in the making and are neither so accurately nor so easily used as the calculating charts to which we shall later come.

We come lastly to the multiple calculating chart, a combination of two or more simple charts of the types described. In the instances considered, the charts have shown only three variables, and therefore been suitable only to equations with two variables beside the root of the equation (in itself a variable but dependent upon the other two variables). More complicated formulae, with three, four, five or more, independent variables, can also be shown by these calculating charts, by the simple trick of joining together a number of individual charts. Thus the first chart can show two inde-



ALIGNMENT CHART FOR SOLUTION OF QUADRATIC AND CUBIC EQUATIONS.

From "Graphical and Mechanical Computation" by Joseph Lipka, published by John Wiley & Sons, by permission.

Fig. 412. A More Complicated Chart for Solving Quadratic and Cubic Equations.

The presence of this chart in this chapter was discovered too late to shift it to its proper place in the chapter on Composite and Zigzag Nomographs, about page 576.

pendent variables on its x and z-scales, and their resultant upon its y-scale. The y-scale of this chart can be used as the x-scale of a second and adjoining chart, a third independent variable appearing on the z-scale of this second chart and the new resultant on its y-scale.

By continuing to join new charts to the old ones, the number of variables in the equation can be increased indefinitely.

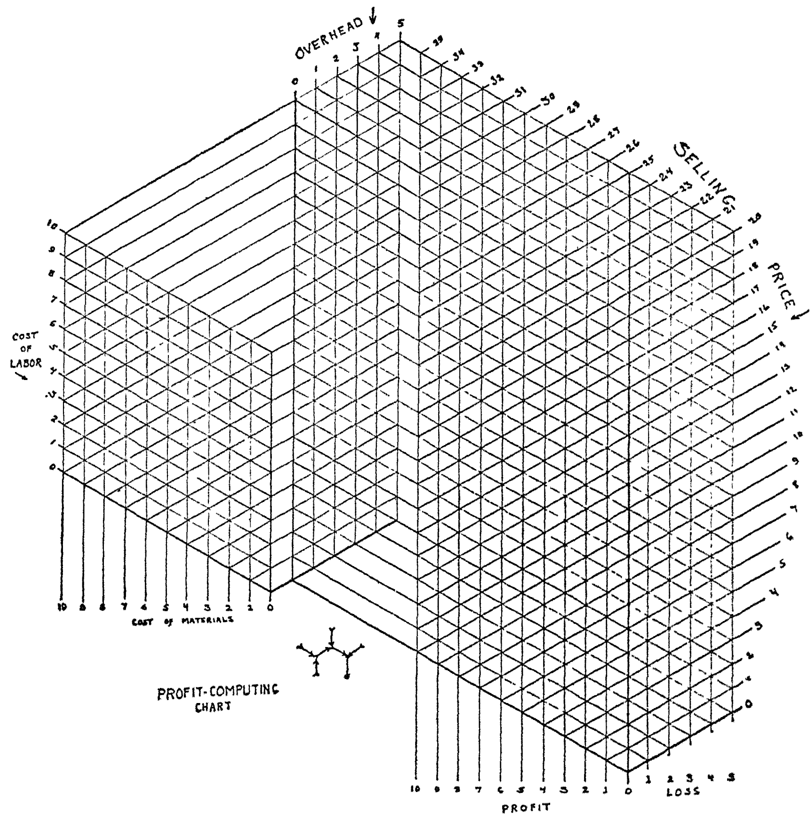
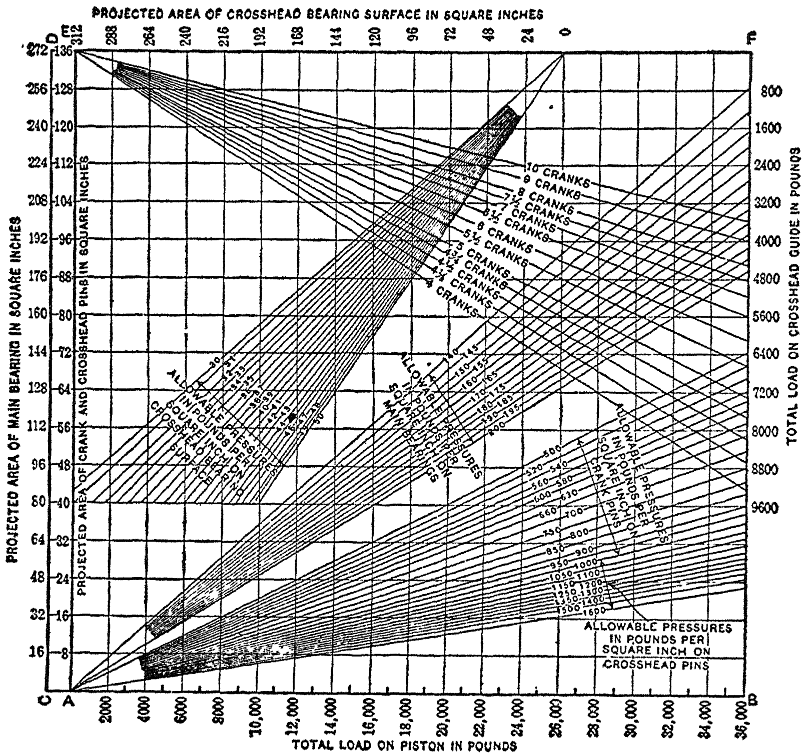


Fig. 413. A Simple Model of Profit-and-Loss Computer.

Where all the independent variables are factorial, we can use the logarithmic projection throughout with parallel straight-line z-curves on every chart, but where some of them are additive, it is not possible to use the logarithmic projection. It is for such cases that the arithmetically projected factorial charts last described come in handy, as by their use the addition processes can be performed upon additive charts and joined

to factorial charts through the common arithmetically projected scales. Often, the various charts are not set side by



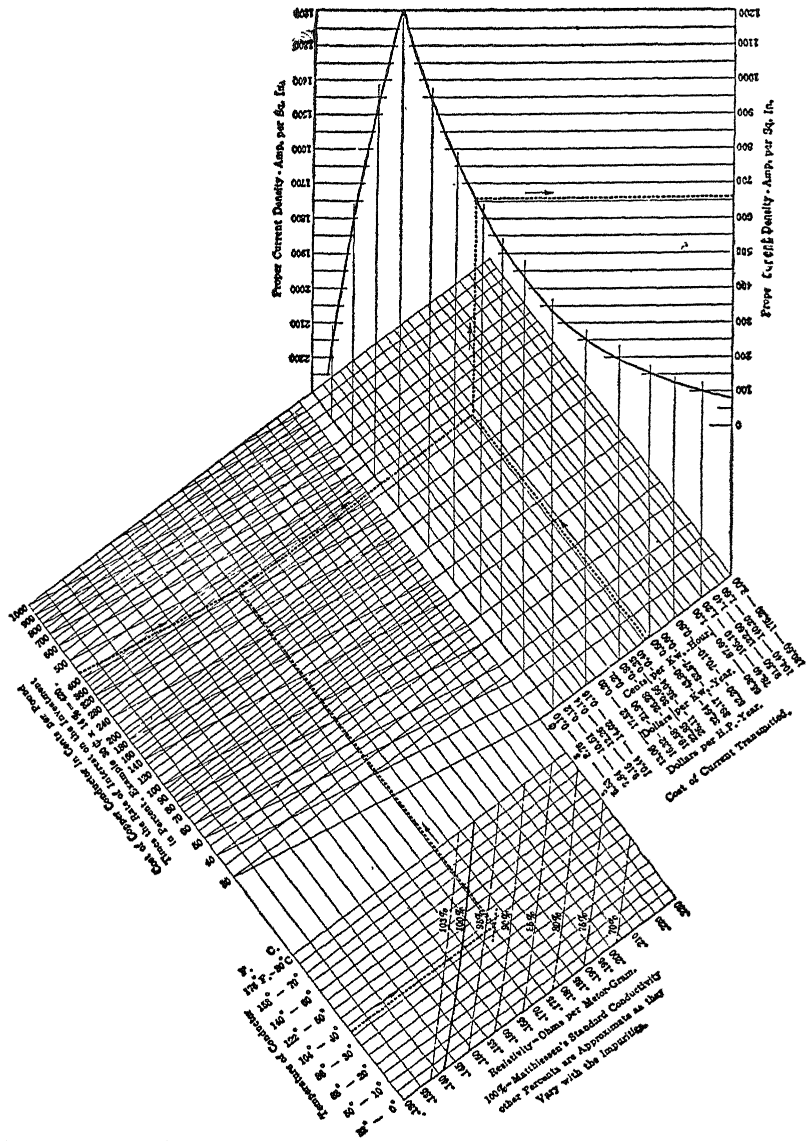
From E. A. Andrews, in "Machinery" and Haskell's "How to Make and Use Graphic Charts."

Fig. 414. A Composite Chart With Many Scales.

Showing Loads on important engine frame members.

side but are superimposed, the reader being asked to follow a sort of mystic maze along these ambiguous co-ordinates till he arrives safely upon the "home"-line scale and meets the answer to his problem awaiting him there.

All of these charts are more sensational then satisfactory. Needless to say, their preparation consumes much time. A large amount of excellent zeal is sometimes displayed by inexperienced chartists in the formation of beautifully-drawn and elaborate chart-forms, suitable for equations with many variables. It is always disappointing to observe the beautiful work and the great energy which has gone into the prepara-



From B. B. Hood, in "Metallurgical and Chemical Engineering," and Haskell's "How to Make and Use Graphic Charts."

Fig. 415. A Simple Combination of Logarithmic and Arithmetical Scales by the Use of a Curve.

Showing the proper current density for copper transmission liner.

tion of these calculating charts, for the truth must be told that both in their preparation and in their reading the multi-

curve equational chart forms are uselessly wasteful of time and energy. Everything that can be accomplished by these elaborate and beautiful charts can be accomplished much more simply, accurately and easily by the use of the charts which will be described in the following chapter, in which the intricate network of co-ordinates[3] and curves alike is entirely omitted and the scales alone are presented upon paper utterly detached from their fields and curves.

---

[3] It is indeed true that the co-ordinates need not be used on the curve if rectangular movable axes (similar to isopleths) on separate transparent sheets be used to project the co-ordinates of the point to the scales where they may be read.

Chapter XLVI

PARALLEL NOMOGRAPHS

In the calculating charts just discussed, we noted that the values which solved a mathematical equation lay along a curve and that the chart was more easily constructed and accurately used when these curves formed straight lines. From this last condition, it is but a step to conclude that the straight-line curves themselves could be omitted and a movable straight-edge (ruler-edge or tightly drawn piece of thread) could be used in their stead, the reader of the chart being required to adjust the straight-edge afresh for each reading. The only objection to this step is that while the lines are straight lines, their angles are arbitrarily set by the problem and the straight-edge must be adjusted at a certain angle or slope before it can be used. But in the charts which we will now consider, this obstruction is removed, the charts being so designed that interpolation by means of the straight-edge is possible in any and every position of the edge, the equation being satisfied always. The straight-line transversal is no longer called a curve, but is now known as an "isopleth," the points through which it passes being always of equal value, that is, forming an equation. The scales are now called "axes" in a wholly isolated sense. The chart itself is called a "nomograph," "nomogram," or "alignment chart," the latter name being obviously derived from the fact that the proper corresponding values are always in perfect alignment.

In the nomograph, the network of co-ordinates and the plotted curves themselves being omitted, there are three scales alone retained. These scales are the scales for the two axes of the curve and the added scale of the diagonal curves themselves. But the scales are so carefully arranged, both as to their projection and as to their position, that the intersection of a straight line or isopleth across two scales always gives the proper corresponding value upon the third scale. The arrange-

ment of these three scales is either parallel or zigzag, so that nomographs can be divided into two classes, the parallel and the zigzag nomographs. Many[1] other forms are possible, but are largely of academic interest. The two principal forms are the simplest and most satisfactory for all work.

The parallel nomographs are based upon the geometrical theorem of similar triangles. If we take the simplest case, in which the three parallel axes or scales, which may be called

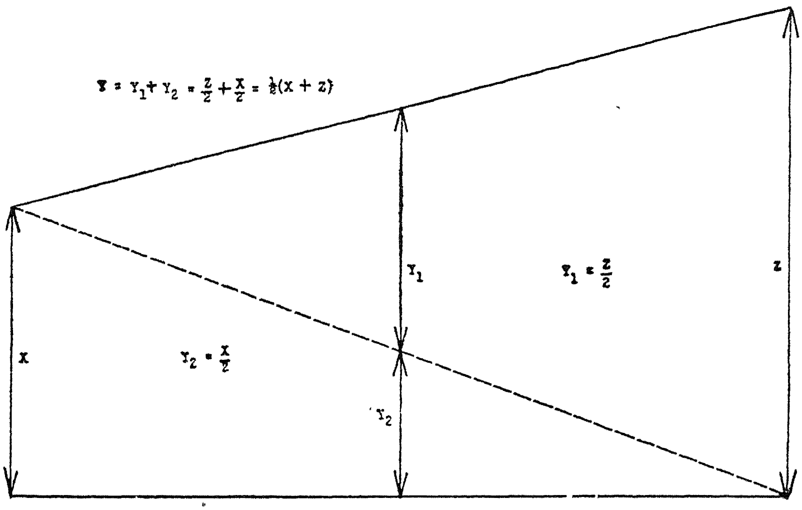


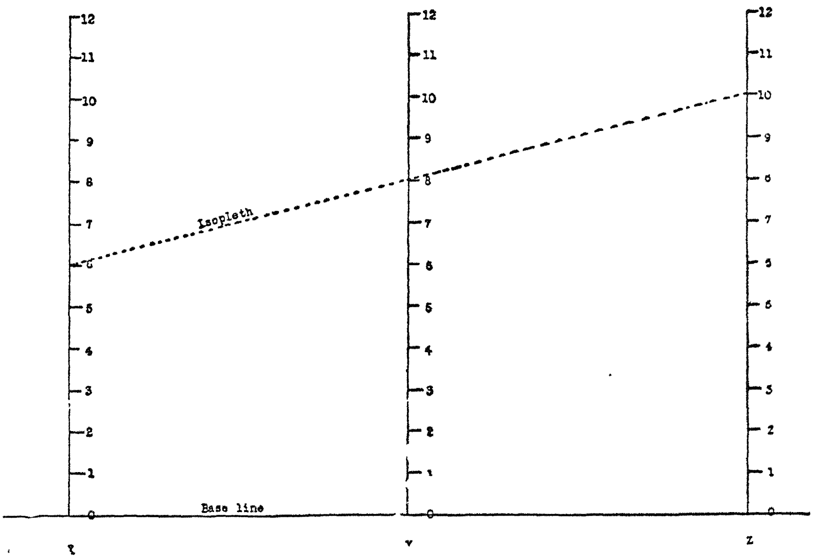
Fig. 416.

the x, y, and z scales, or axes, are so arranged that the two outer ones (let us say the x and z axes) are equidistant from the middle or y-axis, then, we will see that whenever two lines (isopleths) cross these three axes, the distance laid off on the middle axis will be the average (arithmetic mean) of the two distances laid off on the outer axes, between these cross lines.

To make this quite clear let us set three rulers up on end against the wall at equal distances along the wall. With the rulers resting on the floor, their zero-points or lower ends will be in a straight line, the line of the floor itself. Note that the floor here forms an isopleth, the average of the two outer zeros being shown by the middle zero. Now if we hold a piece of string tightly stretched across these rulers, we will

[1] *Cf.* Lipka, Joseph, *Graphical and Mechanical Computation*, John Wiley & Sons. Peddle, John B., *Construction of Graphical Charts*, McGraw-Hill Book Co. and Running, Theodore R., *Empirical Formulas*, John Wiley & Sons.

see that the value on the middle ruler always equals one half the sum of the values on the outer rulers. If the string passes



$y = \frac{1}{2}(x + z)$

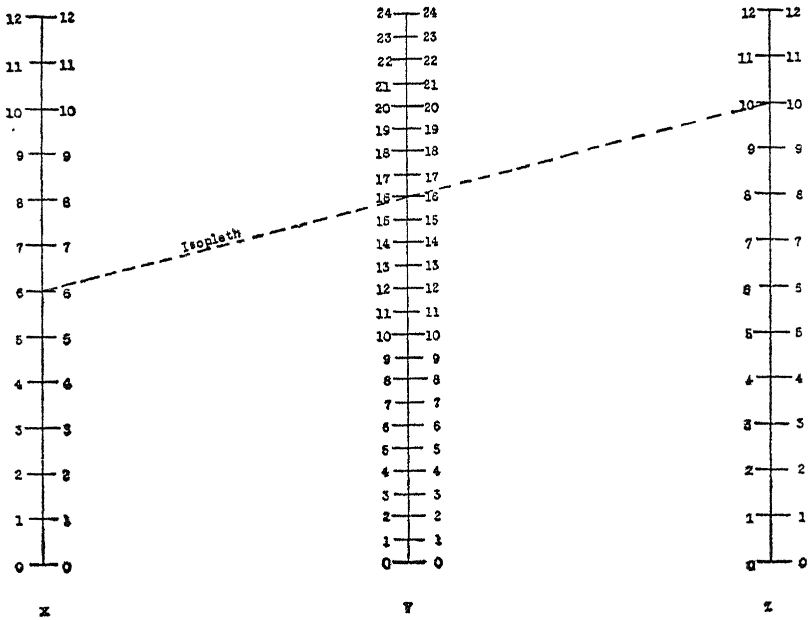
Fig. 417.

across the first or x-rule at the point of six inches, and across the third or z-ruler at the point of ten inches from the floor, it will obviously cross the middle or y-ruler at the point of 8 inches, one half of the sum of six and ten. The general form of the equation is

$$y = \frac{x+z}{2} \text{ or } 2\,y = x + z.$$

To adapt this device to the processes of addition and subtraction is a simple process. Let us merely substitute a ruler calibrated to half-inches for the inch-rule in the middle. In other words, let us substitute for the y-scale a scale with values of Y such that each value of Y is just twice as large a number as its actual y-distance, that is, $Y = 2\,y$. Now the readings Y on the y-scale will be, not the average, but the sum of the readings on the x and z scales. The formula for the chart becomes $2\,y = x + z$, or $Y = X + Z$. And for all positions of the cross-line or isopleth, the intersected points Y, X, and Z, will have the relation $Y = X + Z$. Subtraction may obviously be

performed on such a chart by adjusting the straight-edge or isopleth through any given values on the X and Y or on the



$Y = X + Z$.

Fig. 418.

Z and Y scales, the difference being shown on the remaining (or other outer) scale. For if $Y = X + Z$ then it is clear that $X = Y - Z$ and $Z = Y - X$. The middle or y-axis always carries the minuend in this arrangement, just as it always carried the sum when the same arrangement is used for addition.

Before going further let us again carefully take stock of the algebraic symbols which we shall use in this chapter.[2] The chart, as we have seen, uses one or more sets of three axes, which we shall call the (x), (y), and (z) axes. Along each of these axes we measure distances, x, y, and z, in terms of various units of length, or scale moduli,[3] m_x, m_y, and m_z, all of which are readily convertible into a common unit of length,

[2] See also previous chapter.

[3] Throughout the formulæ for nomographs in this book, the scale-moduli m_x, m_y, and m_z, have been used to indicate the plotting instructions for the variable X, Y, and Z. These formulae, therefore, differ slightly from those of Professor Lipka, in whose book the scale-moduli are used to indicate plotting instructions for the functions of the variables, $f(x)$, $f(y)$, and $f(z)$.

or chart-modulus, m. The values which are entered or calibrated at these distances are X, Y, and Z. These last, X,

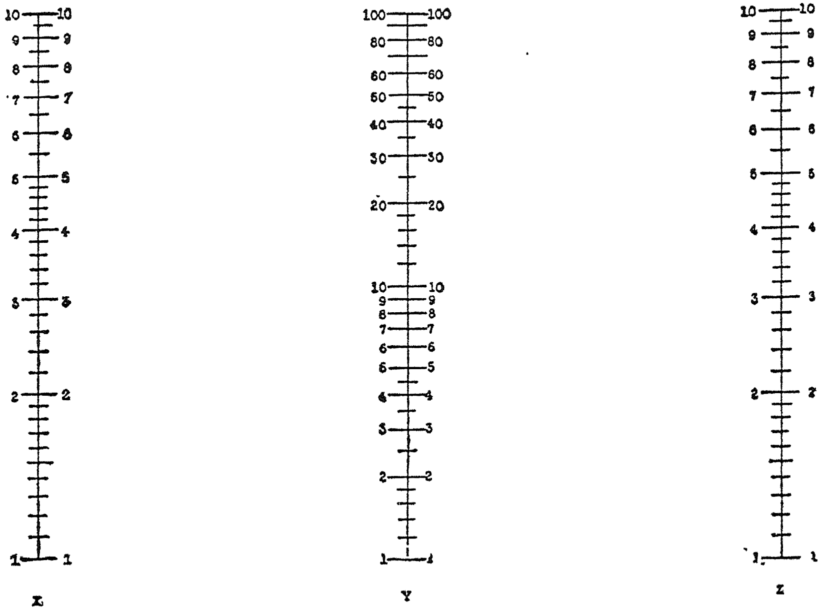


Fig. 419.

Y, and Z, are the symbols of the variables in the equation plotted, they are the scale-figures which appear on the chart, which afford, by their readings along the isopleth, the solutions to the equation.

To adapt this device of three parallel scales to the processes of multiplication and division we need merely change the calibrations on the three scales to a logarithmic projection. As always, the actual distances, x, y, and z, on the three axes, still have the relation $2y = x + z$. But we plot on the x-scale the values of $\log X$, on the z-scale the values of $\log Z$ and on the y-scale the values of $\frac{1}{2} \log Y$, using the plotting equations, $x = m \log X$, $z = m \log Z$, and $2y = m \log Y$ or $y = \frac{m}{2} \log Y$. Since $2y = x + z$, $\log Y = \log X + \log Z$ and $Y = XZ$. The readings of Y are the product of the readings on X and Z, and multiplication is accomplished by adjusting the straight-edge or isopleth through given points on X and Z and reading their

product on Y. Division, like subtraction, is accomplished by adjusting the isopleth through points in the middle and one outer scale, the answer being read on the other outer scale. For if $Y=XZ$, then $X=\frac{Y}{Z}$ and $Z=\frac{Y}{X}$. The middle or y-axis always carries the dividend and the product in this arrangement.

As it will be seen that the distance on the middle scale is always the average of distances on the outer scales, we must expect normally the resultant, that is, the sum or minuend (arithmetically), or the product or dividend (logarithmically) to appear on the central axis only. But a rearrangement of scales can be made when it is desired to place this variable on an outer axis. For this we must use complementary numbers in addition and subtraction, and reciprocals in multiplication and division. That is, we must substitute $0-X$, for X, in addition, and $\frac{1}{X}$ for X or $0-\log X$ for $\log X$ in multiplication. In this arrangement of the additive nomograph (for additions and subtractions) we make $x=m(-X)$ or $-mX$. Then since $2y=x+y$, we have the equation $Y=\frac{2y}{m}=\frac{x}{m}+\frac{y}{m}=-X+Z=Z-X$. And since $Y=Z-X$, then $Z=Y+X$ and $X=Z-Y$. In short by upsetting the scale on the first axis we have exchanged the meanings of the scales on the second and third axes and the third axis now is the resultant (sum or minuend). The rearrangement of the factorial nomograph (for multiplying and dividing) is similar. Here we make $x=m(0-\log X)$, with the result that $-\log X=\frac{x}{m}=\frac{2y}{m}-\frac{z}{m}=\log Y-\log Z$ and $\log X=\log Z-\log Y$ and $X=\frac{Z}{Y}$ or $Z=XY$. The same effect is noticeable as before, the upsetting of one outer scale shifting the meaning of the other two scales, the other outer scale becoming the resultant (product or dividend).

Writers on nomographs are accustomed to attach importance to the position (vertically) of the scales along the axes, a detail to which the cases of reversed or upset scales which we have just considered, naturally leads us. It has been assumed in this discussion that you have kept in mind the idea of three rulers stood up on the floor against the wall, for this makes clear that the three axes must have a common base

line, or isopleth passing through their "zero-distances" (regardless of the calibrations which may be assigned to these dis-

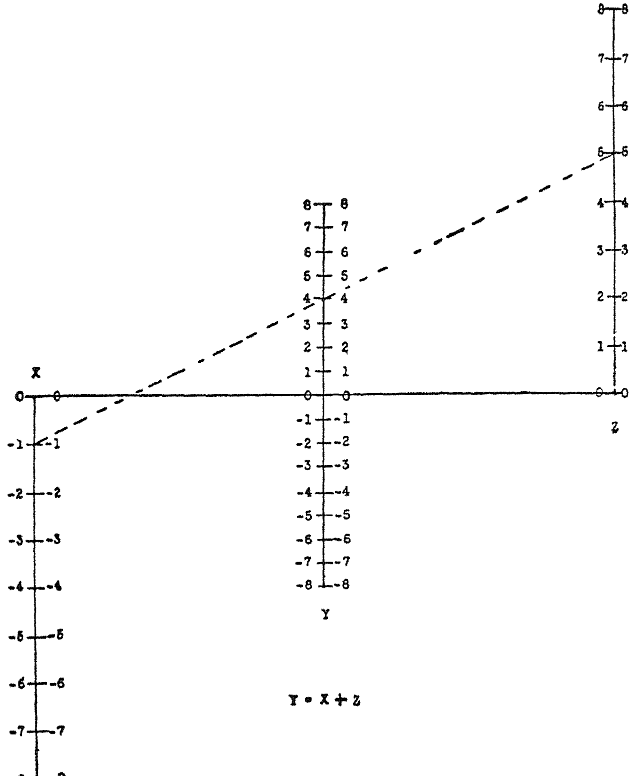
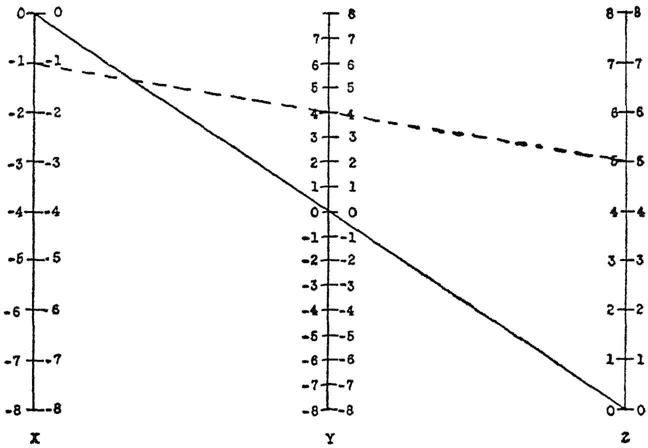


Fig. 420.

tances). Thus when the scales are reversed as in $m(-\log X) = x$ or $\log X = -\dfrac{x}{m}$, it is obvious that we are changing the direction of measurement and counting downward into or below the floor level and an isopleth across the three rulers would have to pass up through the floor. Now it is not at all necessary that the base line (i.e. floor line) be at right angles to the axes (or rulers); our line passing through common "zero-distance" points or "origins" of the axes can be a very steep diagonal. And when one of the scales has been reversed, it is distinctly better to use a diagonal base-line so that the isopleths used in solving the problems by the formula of the chart, shall be as much as possible at right angles to the axes, to facilitate accurate readings.

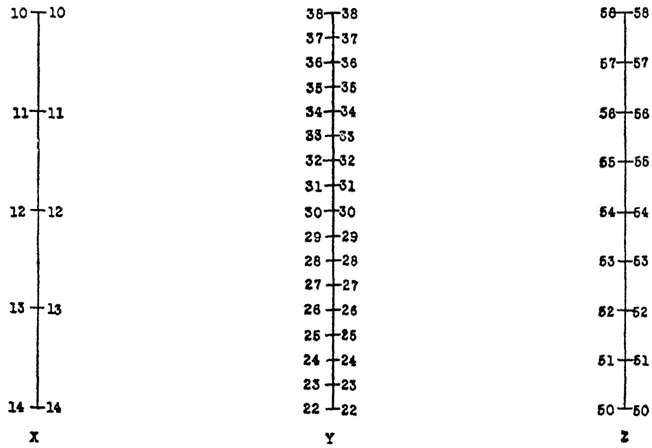
Frequently, in fact, more often than not, the values in which you are interested do not begin at zero, but begin at some distance up the scale, that is to say, the useful or desired range



Y = X + Z

Fig. 421.

of the variables does not come down all the way to the base-line or origin of the axes. In that case again, it is well to use a



Y = Z - 2X

Fig. 422.

diagonal base-line, in order that you may omit the lower parts of the scale entirely, together with the base-line itself, on your

finished chart. Another way to achieve the same result is to alter the calibrations alone, so that $x = m(X + a)$, $y = \frac{1}{2}m(Y + b)$ and $z = m(Z + c)$, (in which a, b, and c are constants which in themselves satisfy the formula). In general, the values of these constants should be such that they are equivalent to the lower limits of the desired ranges of the variables. The real object of the diagonal base-line or diagonal zero isopleth, is to make all isopleths which will be used on the chart as perpendicular to the scales as possible. The nearer to a right-angle the intersection of isopleth and axis becomes, the more sharply the two lines cut each other and the more easily will accurate readings be made.

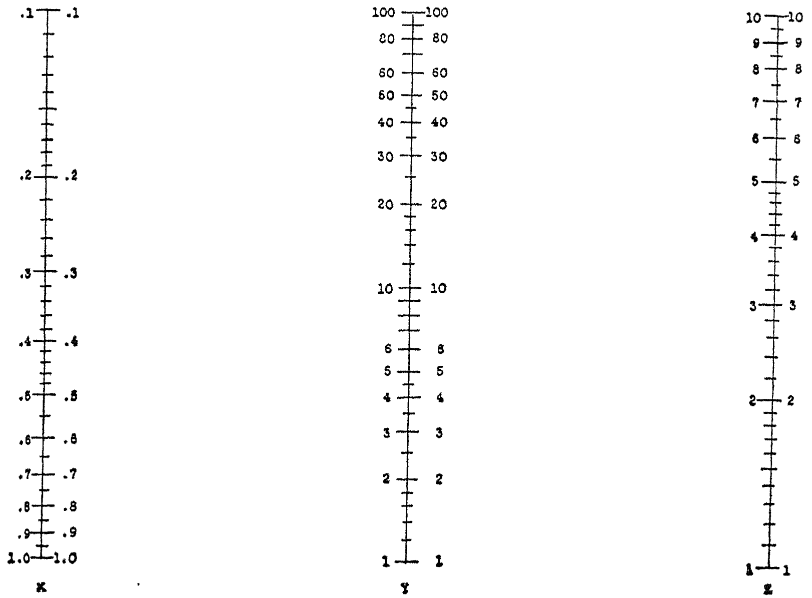
This brings us to the important element of the range of the variables. For it is not necessary, nor even possible, to picture all the possible values of a variable upon a chart. In actual problems the independent variables will usually be found to fluctuate between certain limits. It is thus unnecessary to use a scale so great that it shows values in excess of the maximum limits, or to include on the scale the values below the minimum limits. Space is conserved and detail gained by making the range of the scale conform to the range of the useful values of the variable. And when the ranges of each of the two independent variables have been set, it is easy to find the range of the resultant or dependent variable (the root of the equation). In the previous chapter we have indicated a method of noting these limits, thus

$$Y = X \left| \begin{matrix} 10 \\ 0 \end{matrix} \right| + Z \left| \begin{matrix} 5 \\ 0 \end{matrix} \right|, \text{ or } Y \left| \begin{matrix} 15 \\ 0 \end{matrix} \right| = X \left| \begin{matrix} 10 \\ 0 \end{matrix} \right| + Z \left| \begin{matrix} 5 \\ 0 \end{matrix} \right|.$$

A variety of methods are at hand for confining the chart to these ranges. We may place the lower limits at the zero distances or origins of each axis. Or we may place the maxima upon the level (or base-line). Or best of all, we may place the mid-points along each range upon a level isopleth. The advantage of the last method is that all the possible isopleths will then cross the axes at angles nearer to a right angle than by any other arrangement. Having approximately positioned our scales with this object in view we do not actually need to calculate the values of the mid-points (fractional as these may be), for plotting; we need only calculate the values for any round numbers and precisely position the scales about them.

An important point in the making of the chart is its total size and proportions. Both its height and width should be great enough to serve whatever purposes of convenience in use, legibility and detail of readings, visibility at certain distances, or success in reproduction and reduction, will naturally obtain in chart-making, but the width should always be at least as great and if possible half as great again as the height. If the chart is too narrow many of its useful isopleths will cross the axes at such small angles that correct readings are difficult. If the chart is too wide, the isopleths will all cross at very good angles but the scales will be so closely compressed as to make detailed readings hard. The best form in general is one in which the most steeply sloping isopleths cannot cross the axes at smaller angles than from 45 degrees to 60 degrees. The height should be approximately two-thirds the width.

The final consideration is the choice of axis for the dependent variable. By the dependent variable is meant the variable whose values are sought from given values of the



$Y = \frac{Z}{X}$

Fig. 423. The Inverted X-Scale.

other variables. Of course it often happens that the same equation is often used backwards, and that at times one variable is sought from given values of the other two and at times another is sought. But usually there is one variable which is most likely to be the unknown and this should be treated as the dependent variable. The best axis for the dependent variable is always, *ceteris equibus*, the (y) axis. For then all needed isopleths will lie within the limits of the two outer scales and the three scales can be of roughly uniform height.

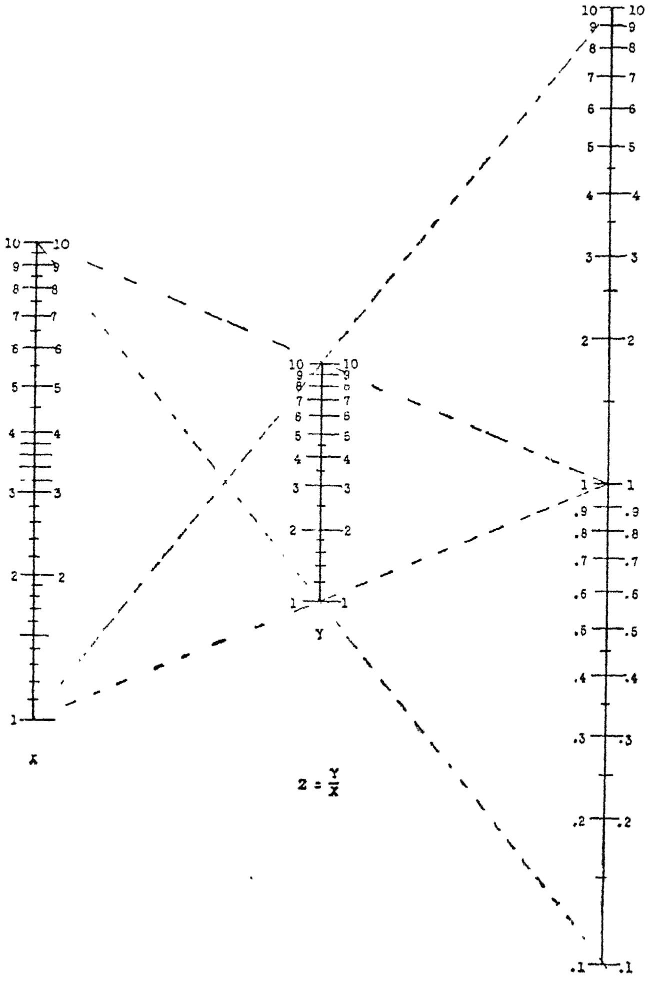


Fig. 424. The Use of an Outer Scale for the Unknown Variable is Not Good.

Were the known variable placed upon an outer axis, it is clear that it would have to be extending above and below the levels of the other axes unil it included the most extremely sloping isopleths which could be drawn through the central and other outer scales. The result would be a chart of very irregular appearance, wasteful in space and involving less accurate readings because of smaller angles of intersection between isopleth and axis. The danger of errors in placing the isopleth would be four times as great, since the errors in positioning the known values may be doubled upon the unknown scale, whereas they are halved when the unknown scale is on the central axis.

It has been the purpose of the foregoing discussion to be suggestive rather than definitive, of the general principles of the parallel nomograph. It remains to examine this chart analytically. This will lead us at once to a generalized form of the parallel nomograph, with important modifications which make it far more flexible, in use. The chart has so far been considered only with equidistant axes.

The geometrical proposition of similar triangles can equally be applied to axes which are not equidistant. When the interval between axes (x) and (y) is equal to that between (y) and (z), then the formula for actual distances is, as we have seen, $y=\frac{z}{2}+\frac{x}{2}$ or $2\ y=x+z$. And if we denote the total distance

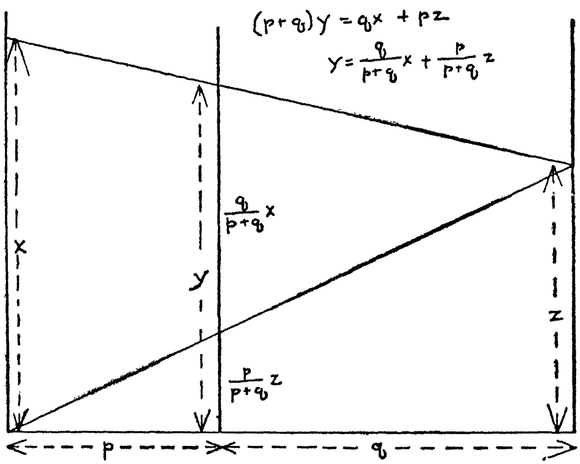


Fig. 425.

between the (x) and (z) axes, either measured perpendicularly to the axes or along the base-line or along any isopleth, as "$p+q$," taking "p" as the part between the (x) and (y) axes and "q" as the part between the (y) and (z) axes, we may write the formula for the distances or axes cut between isopleths as

$$y = \frac{(q)}{(p+q)}x + \frac{(p)}{(p+q)}z = \frac{qx+pz}{p+q}, \text{ or } (p+q)\,y = qx + pz.$$

This formula is applicable to any parallel nomograph, no matter at what distances from each other the axes may be placed. And the significance of "p" and "q" are very easily seen. They are the coefficients of the x and z variables in the additive formula $2\,Y = X + Z$, and the corresponding exponents in the factorial formula $Y^2 = XZ$, becoming coefficients in the corresponding equation $2 \log Y = \log X + \log Z$. In short the complete formulae are, for additive charts $(p+q)\,y = qx + pz$, and for factorial charts, $y^{(p+q)} = x^q\,z^p$. Obviously when p and q are equal they may be written as 1 so that we have $2\,y = 1\,x + 1\,z$ and $y^2 = x^1\,z^1$. So that the formula at once explains the half-size scales taken for the middle axes. Also when we reversed or upset one scale, we were in the additive formula inserting a -1 coefficient, making the value of $q = -1$. We were then obliged to shift the other outer axis into position midway between the first and second axes (a process which we spoke of as exchanging meanings of scales) so that p became $+2$ and $(p+q)$ became $+1$, so that the formula[4] became $1\,y = -1\,x + 2\,z$. Thus the general formula $(p+q)\,y = qx + pz$ covers all cases of the parallel nomographs.

We are now ready to lay down the rules for the construction of the parallel nomograph. In the first place, we have an

---

[4] Or, calling the y-axis z, because it is now the third, and the z-axis y, because it is now the second, we have $+1z = -1x + 2y$, which agrees exactly with the formula for reversed scales. So doing, we maintain the symbols (x), (y) and (z) for the axes strictly in the order in which these axes appear on the chart.

It is obviously better to permit the symbols (x), (y), and (z) to adhere to the axes wherever they appear, regardless of their order upon the page, as the general formulæ then applies consistently and without confusion. In the text from this point on, this has been done, and the (x)-axis need no longer be the first, the (y)-axis the second, nor the (z)-axis the third; but the algebraic signs of p and q will signify changes in position, and the algebraic signs of the scale-moduli will be significant of the direction of plotting.

equation of the general type $Y = AX\left|\frac{H_x}{L_x}\right| + CZ\left|\frac{H_z}{L_z}\right| + K$, which[5] we wish to present upon a chart or diagram which has the relations of $qx + pz = (p + q)y$. We give this diagram any height, Tm, we wish and approximately half as much more width. If, as is most convenient, we let the chart-modulus, m, equal 1 inch, then T is the total height of the chart, or length of each scale, in inches. Now along these scales we propose to plot the values of the independent variables, X and Z, from their lowest, L, to their highest, H, useful values. Call the difference between these extremes the range, R, of the variable, then

$$R_x = H_x - L_x \qquad R_z = H_z - L_z$$

We can easily plot the values X and Y through these ranges in these given lengths by the method of triangulation if, as is generally the case, the intervals are not even fractions of the inch.[6] Then draw an isopleth through any convenient values of X and Z on the (x) and (z) scales and we know that the corresponding value of Y lies somewhere along this isopleth. Substitute these values of X and Z in the equation and learn the corresponding value of Y. Select a second convenient value for X and substituting it and the Y-value for X and Y in the equation, solve and get a second corresponding value of Z. Draw a second isopleth through the second values for X and Z on the (x) and (z) scales and since the value of Y has remained unchanged, we know that the (y) axis passes through the intersection of the two trial isopleths, and is parallel to the other axes. Now solve a few more equations containing convenient values of X and Y and draw their isopleths and you will rapidly calibrate the (y) axis with its Y values. After a few points have been plotted the rest of the Y-scale can be put in by a ruler, and the method of triangulation. If these directions are carried out the entire chart will be finished in a short time.

The student will look however, for an analytical method which will define mathematically the various scales and their

[5] Within the short vertical parallel lines in the equation are inserted the high, H, and low, L, values of the variables which will be required. These maxima and minima of the ranges are merely memoranda which do not affect the equation in the least, and can be omitted from the equation and noted elsewhere, if they render the equation confusing.

[6] For the precise adjustment of scales to given sizes by the method of triangulation, see Chapter XVII, page 185.

PARALLEL NOMOGRAPH:

EQUATION:

$$Y = AX + CZ + K$$

$$2u = 3v + 5w$$

ASSIGNMENT OF SYMBOLS:

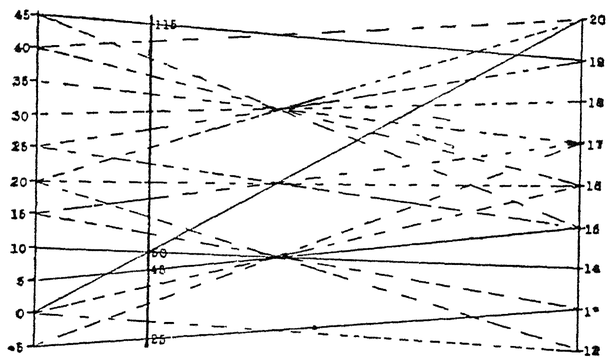
TO INDEPENDENT VARIABLES: LET $X = v$

$Z = w$

TO DEPENDENT VARIABLE: $Y = u$

LIMITS OF USEFUL VARIATIONS:

MAXIMUM	$H_x = 45$	$H_z = 20$
MINIMUM	$L_x = -5$	$L_z = 12$
RANGE	$R_x = H_x - L_x = 50$	$R_z = H_z - L_z = 8$



CALCULATIONS FOR Y-SCALE:

WHEN $X = 0$ AND $Z = 20$; $Y = (3/2)0 + (5/2)20 = 50$

$Y = 50$ $Z = 14$; $X = (2/3)50 - (5/3)14 = 10$

$Y = 45$ $Z = 15$; $X = (2/3)45 - (5/3)15 = 5$

$Y = 25$ $Z = 13$; $X = (2/3)25 - (5/3)13 = -5$

$Y = 115$ $Z = 19$; $X = (2/3)115 - (5/3)19 = 45$

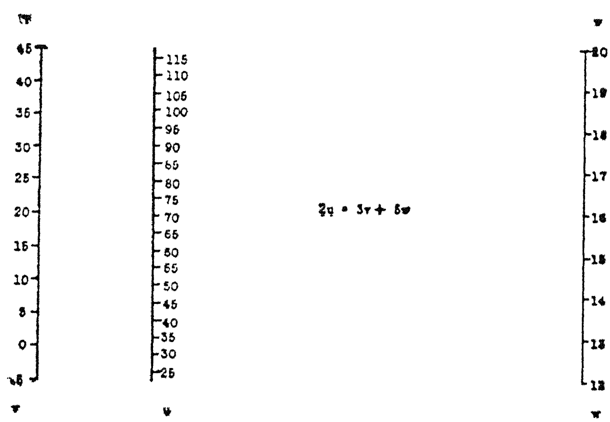


Fig. 426. Construction of the Parallel Nomograph—I.
Finding the unknown scale by trial isopleths.

positions. For this we must know the scale-moduli, m_x, and m_z, as distinct from the chart-modulus, m. The scale-modulus is the interval or unit distance along the scale between the unit values of the variables, X or Z, and is obviously written

$$m_x = \frac{T}{R_x} m \qquad\qquad m_z = \frac{T}{R_z} m$$

or letting $m = 1$ inch)

$$m_x = \frac{T}{R_x} \text{ inches} \qquad\qquad m_z = \frac{T}{R_z} \text{ inches}$$

Now in plotting it is always a convenience to adopt units of length such that they can be laid off from an engineer's hexagonal rule and do not have to be specially projected by triangulation. The engineer's rule divides the inch into 1, 2, 3, 4, 5, or 6 parts and decimals multiple or submultiple thereof, such as 10, 20, 400, 3000, etc. If we describe the scale moduli by the numbers of them which go to make up the chart modulus, m, and call this number S, then

$$S_x m_x = m \qquad\qquad S_z m_z = m$$

and

$$m_x = \frac{m}{S_x} \qquad\qquad m_z = \frac{m}{S_z}$$

or, letting $m = 1$, (that is, one inch)

$$m_x = \frac{1}{S_x} \qquad\qquad m_z = \frac{1}{S_z}$$

Thus as we have previously seen, the reciprocal of the scale modulus, when $m = 1$ inch, is the number of intervals per inch and serves as an index of the proper side of the engineer's rule to use in plotting. Combining the two equations for m_x we can eliminate it and keep measurements in terms of its reciprocal, S, thus

$$\frac{m}{S_x} = \frac{T}{R_x} m \qquad\qquad \frac{m}{S_z} = \frac{T}{R_z} m$$

or

$$S_x = \frac{R_x}{T} \qquad\qquad S_z = \frac{R_z}{T}$$

If S in these equations becomes either 1, 2, 3, 4, 5, or 6, or any multiple of any power of ten, we can of course

plot the (x) and (y) scales directly from the engineer's rule. If it does not do so at once (and it usually doesn't) always make it do so by altering the values of R and T slightly; that is, change the length of the scale, T, or increase its range, R, or do both. Slight alterations of this kind do not appreciably affect the size or usefulness of the chart. We shall write these altered values in small letters, thus

$$S_x = \frac{r_x}{t_x} \qquad\qquad S_z = \frac{r_z}{t_z}$$

So we see that by selecting values for r and t which are close to the original values R and T, but which make S precisely indicate a side of the engineer's rule, we make the chart even simpler to draw.

Now we have seen that the chart has the linear relations, $qx + pz = (p+q)y$. If the chart is to express the equation, $Y = AX + CZ + K$, and we decide (since it is easiest) to correct for the added constant K along the (y) scale, then we write

$$AX + CZ = Y - K$$

and the chart must have the values

$$qx = m\ (AX) \qquad pz = m\ (CZ) \quad (p+q)\ y = m\ (Y - K)$$

$$x = \frac{A}{q}\,mX \qquad z = \frac{C}{p}\,mZ \qquad y = \frac{1}{p+q}\,m\ (Y - K)$$

In this the distances, x, y, and z, are taken from origins which will be real if L_x, L_z, and L_y are zero, but imaginary if L_x, L_z, and L_y are other than zero and the zero or base-line is not shown on the chart. Also we know that by definition

$$x = m_x X \qquad z = m_z Z \qquad y = m_y\,(Y - K)$$

Hence

$$m_x = \frac{A}{q}\,m \qquad m_z = \frac{C}{p}\,m \qquad m_y = \frac{1}{p+q}\,m$$

But from above

$$m_x = \frac{m}{S_x} \qquad\qquad m_z = \frac{m}{S_z}$$

Hence

$$\frac{1}{S_x} = \frac{A}{q} \qquad\qquad \frac{1}{S_z} = \frac{C}{p}$$

$$q = AS_x \qquad\qquad p = CS_z$$

PARALLEL NOMOGRAPH: EQUATION:

$$Y = AX + CZ + K$$
$$2u = 3v \quad + 5w$$

SYMBOLS, LIMITS, AND DEPENDENCE:

$$AX\begin{vmatrix} H_x \\ \\ L_x \end{vmatrix} \quad + CZ\begin{vmatrix} H_z \\ \\ L_z \end{vmatrix} \qquad = Y - K$$

$$\frac{3}{2}X\begin{vmatrix} 45 \\ \\ -5 \end{vmatrix} \quad + \frac{5}{2}Z\begin{vmatrix} 20 \\ \\ 12 \end{vmatrix} \qquad = Y - 0$$

SIZE OF CHART (HEIGHT): LET T = 4 inches

RANGE OF VARIATION:

$R_x = H_x - L_x = 50$ $\qquad R_z = H_z - L_z = 8$

RULERS FOR PLOTTING:

$S_x = r_x \,/\, t_x$

50 / 4 = 12.5

50 / 2.5 = 20

50 / 5 = $\underline{10}$

$S_z = r_z \,/\, t_z$

8 / 4 = $\underline{2}$

INTER-AXIAL DISTANCES:

$q = AS_x = (3/2)\ 10 = 15$ $\qquad p = CS_z = (5/2)\ 2 = 5$

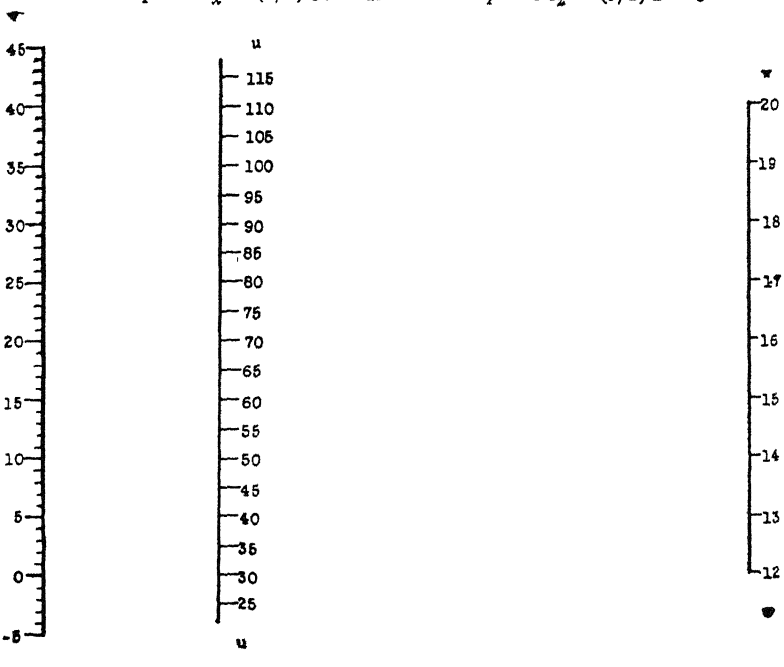


Fig. 427. Construction of the Parallel Nomograph—II.
Finding rulers with which to plot the known variables and a formula with which to position the unknown axis. (The ruler values adopted are underlined in the worksheet above.) The unknown variable is still plotted by trial isopleths.

Here we have a convenient formula for the distances, p and q, between the three axes, (x), (y), and (z). Having found from the range and total length of the scales the convenient sides of the rule to use in plotting them $(S=\frac{r}{t})$ we merely multiply these by the coefficients, A and C, of the variables in the equation to get the horizontal distances between axes. You will notice that the chart-modulus, m, has cancelled out of the equations, so that p and q can be measured in any units which will make their sum be the desired width of the chart. These devices have obviated the first two trial solutions of the formula for the purpose of locating the (y) axis.

Lastly we come to the formula for the plotting of the (y) scale itself. Just as we wrote $m_x=\frac{m}{S_x}$ and $m_z=\frac{m}{S_z}$, so we can write $m_y=\frac{m}{S_y}$ and our object will be to find S_y, such that it too can be plotted directly from an engineer's rule. Above, we see that $m_y=\frac{1}{p+q}m$, hence[7]

$$\frac{1}{S_y}=\frac{1}{p+q} \qquad\qquad S_y=p+q=AS_x+CS_z$$

A and C, of course, are fixed, and you will often find it impossible to so adjust S_x and S_z that while they indicate sides of the engineer's rule, S_y does the same. It is generally necessary, to go back to the original elements of S_x and S_z, namely $\frac{r_x}{t_x}$ and $\frac{r_z}{t_z}$, and alter one or both of them until the desired result is achieved. A convenient plan is to set down columns for each of the elements, t_x, r_x, S_x, q, S_y, p, S_z, r_z, and t_z, so that you can try a number of different values of T and R, for each variable before you give up hope.

When S_y cannot be made to conform to a side of the rule, it is still always easy to project specially by the method of

[7] Since obviously $m_y=\frac{m}{S_y}$

PARALLEL NOMOGRAPH: EQUATION: $2u = 3v + 5w$

$$AX\left|\begin{matrix}H_x\\ \\ L_x\end{matrix}\right| + CZ\left|\begin{matrix}H_z\\ \\ L_z\end{matrix}\right| = Y - K$$

$$\frac{3}{2}X\left|\begin{matrix}45\\ \\ -5\end{matrix}\right| + \frac{5}{2}Z\left|\begin{matrix}20\\ \\ 12\end{matrix}\right| = Y - 0.$$

PLOTTING INSTRUCTIONS:

t_X	r_X	S_X	q	S_Y	p	S_Z	r_Z	t_Z
.A		$\frac{r_X}{t_X}$	AS_X	$q+p$	CS_Z	$\frac{r_Z}{t_Z}$		
T=6	R_X=50		$A=\frac{3}{2}$		$C=\frac{5}{2}$		R_Z=8	T=6
6	50	$8\frac{1}{3}$	$12\frac{1}{2}$	$15\frac{5}{6}$	$3\frac{1}{3}$	$1\frac{1}{3}$	8	6
5	50	$\underline{10}$	15	$\underline{20}$	5	$\underline{2}$	8	4
2.5	50	20	30	37.5	7.5	8	9	3
2.5	50	$\underline{20}$	30	$\underline{40}$	10	$\underline{4}$	8	2
8.5	51	6	9	11.5	2.5	1	8	8
10	50	5	7.5	9	1.5	.6	9	15
10	50	$\underline{5}$	7.5	$\underline{10}$	2.5	$\underline{1}$	8	8

Fig. 428. Construction of the Parallel Nomograph—III.
Finding ruler-values with which to plot the unknown variable. (Any of the sets of underlined rulers for the three variables can be used, according to the size, Tx and Tz, which is desired for the chart.) Worksheet only is shown here.

triangulation. There is no more need to make trial isopleths for specially computed values of the variables, and we would only make two or three of these when the chart is completed to check it up for accuracy. The whole problem of charting the equation $Y = AX\left|\begin{matrix}H_x\\ L_x\end{matrix}\right| + CZ\left|\begin{matrix}H_z\\ L_z\end{matrix}\right| + K$ is reduced to the following simple steps:

$$AX\left|\begin{matrix}H_x\\ L_x\end{matrix}\right| + CZ\left|\begin{matrix}H_z\\ L_z\end{matrix}\right| = Y - K$$

$S_x = \dfrac{r_x}{t_x}$ when r_x approximates $R_x = H_x - L_x$.

$S_z = \dfrac{r_z}{t_z}$ when r_z approximates $R_z = H_z - L_z$.

$p = CS_z$, $\quad q = AS_x$, and $S_y = CS_z + AS_x$.

We know T, the height which we wish to give the chart and which t_x and t_z approximate, and we know the width which we wish to give the chart, approximately $\dfrac{3T}{2}$, which will be divided by the axes in the ratio of p and q. It only remains for us to select mid-points in the ranges of X and Z and place them, with the corresponding value of Y, upon a common horizontal isopleth and plot the scales about them.

When the equation is factorial instead of additive, it has the form

$$Y = \left(X \left| \begin{matrix} H^x \\ L_x \end{matrix} \right| \right)^A \left(Z \left| \begin{matrix} H_z \\ L_z \end{matrix} \right| \right)^C K$$

or

$$\log Y = A \log X \left| \begin{matrix} H_{\log x} \\ L_{\log x} \end{matrix} \right| + C \log Z \left| \begin{matrix} H_{\log z} \\ L_{\log z} \end{matrix} \right| + \log K$$

and the same treatment may be followed precisely. It is more convenient, however, to plot directly from a log rule, if one is handy, than from an engineer's rule and a table of logs. We therefore drop the engineer's rule and use the calibrations on a slide-rule, if one is available. The simpler slide-rules have two scales, one a single and the other a double deck, in the length of 25 centimeters. Better slide-rules have also a three-deck scale. Taking the modulus, m, of the chart as 25 centimeters instead of 1 inch (it does not make any difference in the planning equations just listed since m has been cancelled out of them) we now measure T, t^x, and t^z, in units of 25 centimeters, roughly 10 inches, and take $S = 1$ to indicate the single, $S = 2$ the double, and $S = 3$ the triple deck scales. In short, when S_x, S_z, or S_y can be made equal to 1, 2, or 3, we can plot scales directly from the slide-rule.

Complicated formulae cannot often be made to yield direct ruler-copying values of all three, S_x, S_z, and S_y, at the same time, either for the additive or the factorial charts. This difficulty is most frequently encountered in the factorial charts because of the more limited number of different rulers. When this is the case, the method of parallel triangulation can, of course, be used for all other values of S_x, S_z, and S_y. But most convenient of all is a set of radiating triangulation

sheets, such as are included in Professor Lipka's book,[8] which can be folded at any value of m and will give all possible projections of the arithmetic or logarithmic scales. When such devices are used the chart-maker has no occasion to seek certain values of S, but can work with any scale-moduli whatever, so that his equations become (when m, the chart-modulus, is 1 inch):[9]

$$m_x = \frac{T}{R_x} \qquad m_z = \frac{T}{R_z}$$

$$q = \frac{A}{m_x} \qquad p = \frac{C}{m_z}$$

$$m_y = \frac{1}{p+q}$$

and he can plot directly from his sheet of scales, folded at the proper scale-modulus.

The general equations which have been given, namely,

$$Y = AX\left|\frac{H_x}{L_x}\right| + CZ\left|\frac{H_z}{L_z}\right| + K$$

$$\text{and } Y = \left(X\left|\frac{H_x}{L_x}\right|\right)^A \left(Z\left|\frac{H_z}{L_z}\right|\right)^C K$$

are usually found in simplified form, K being 0 in the additive (first) form, or 1 in the factorial (second) form. When a coefficient (in the additive) or exponent (in the factorial) is attached to the dependent variable, Y, it can be transferred to the other variables so as to clear Y, by division or involution. When the signs of either X or Z are negative, the sign must be treated as part of the coefficient and so transferred to the scale-modulus or ruler-index (S), to indicate that all values are plotted downward instead of upward.

Variations of this parallel nomograph will occur to the student, such as charts for the equation

$$Y = X^A Z^C K + D,$$

which must be turned into

$$\log (Y - D) - \log K = A \log X + C \log Z$$

[8] To the chart-maker, Professor Lipka's book *Graphical and Mechanical Computation* is well worth its cost, if for no other reason than for the useful scales it contains in a pocket in the rear cover of the book. These scales carry radiating lines from a common center to all parts of a ten-inch uniform and a ten-inch logarithmic (single-deck) scale. By folding these sheets appropriately, these scales can be obtained from the radiating lines at any desired smaller scale. They amount to complete outfits for the triangulation method of scale adjustment.

[9] As is the case with Lipka's chart-formulae.

PARALLEL NOMOGRAPH: EQUATION: $27.2\,W = 13.7\,R^2\,11\,g^3$

$$A X \Big|_{L_x}^{H_x} + C Z \Big|_{L_z}^{H_z} = Y - K$$

$$2 \log R \Big|_{.25}^{25} + 3 \log g \Big|_{.05}^{50} = \log W - \log \frac{13.7 \times 11}{27.2}$$

$$= \log W - \log 5.17$$

CHARTING INSTRUCTIONS:

T = 6 inches

$R_X = \log 25 - \log .25 = \log 100 = 2$ $\quad$ $R_Z = \log 50 - \log .05 = \log 1000 = 3$

$m_X = T/R_X = 6/2 = 3$ inches $\quad$ $m_Z = T/R_Z = 6/3 = 2$ inches

$q = A/m_x = 2/3 = .667$ $\quad$ $p = C/m_Z = 3/2 = 1.5$

$$m_Y = 1/(p+q) = \frac{6}{13} = .461 \text{ inches}$$

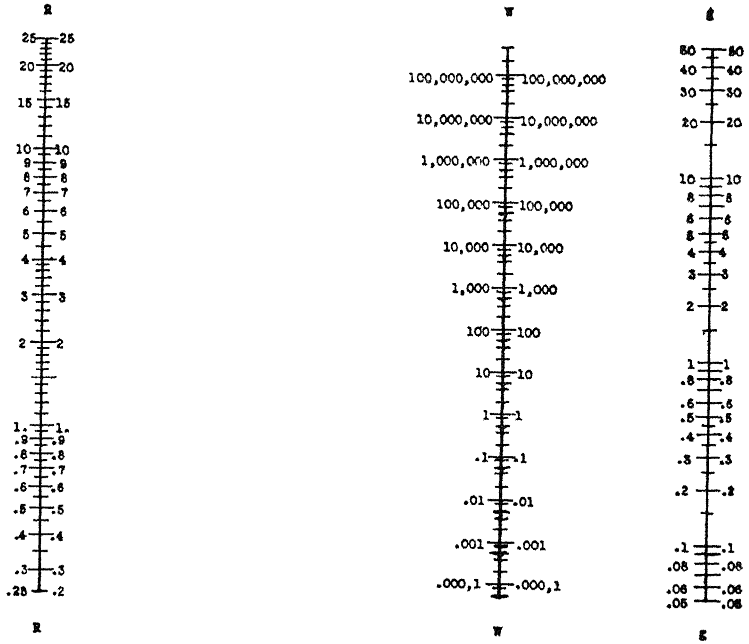


Fig. 429. Construction of the Factorial Parallel Nomograph.
Finding slide-ruler plotting scales for all variables and locating the unknown variable axis, all by formula. (Only the adopted values are shown in the work-sheet, but a columnar form similar to that used in the last figure, is useful to compare different values before selection.)

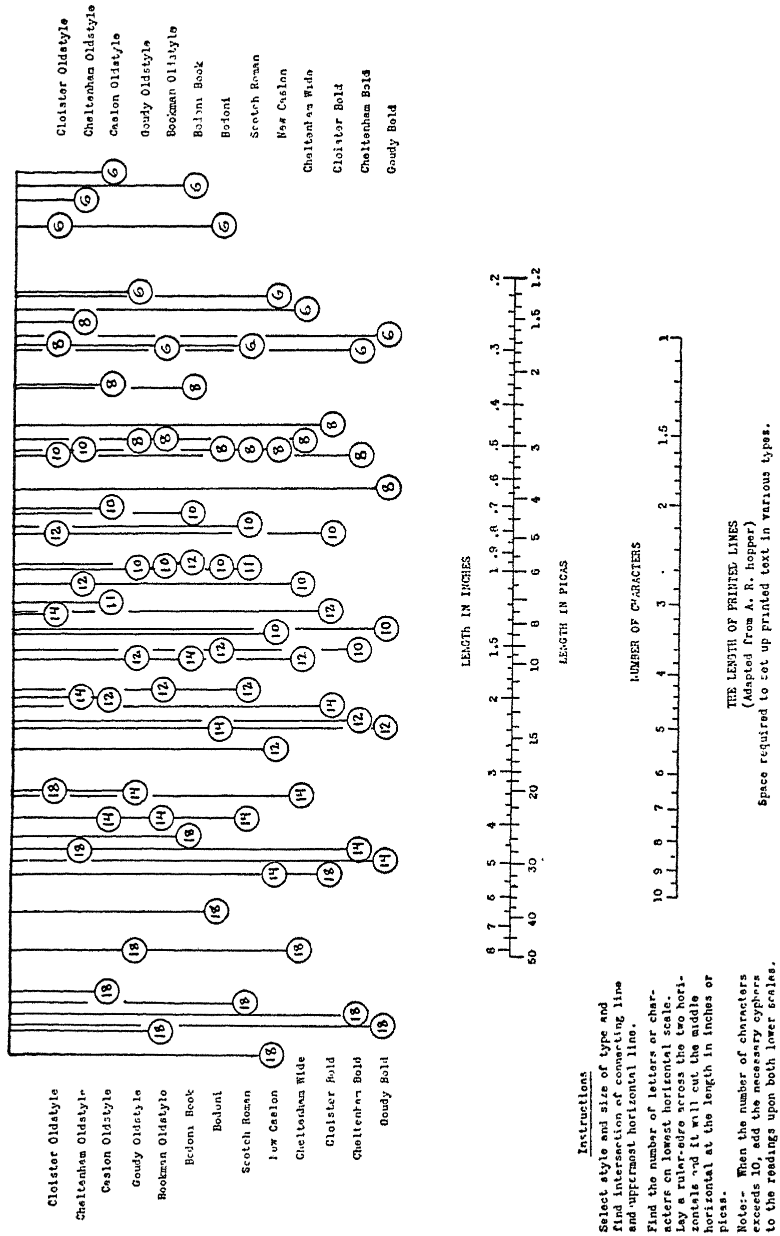


Fig. 430. Chart for Determining Size of Type.

Showing the space required to set up manuscript in specified sizes and styles of type. The latter have been indicated on the chart on a properly graduated dummy axis, but without numerical calibrations, of characters per inch, as these would be useless.

and involve therefore a special logarithmic projection with shifted zero. The exponential equation

$$Y = A^X C^Z K$$

can be turned into

$$\log Y - \log K = (\log A)\, X + (\log C)\, Z$$

and involves a mixture of log and arithmetical scales, the scales of (x) and (z) being arithmetical. So does the equation

$$Y = A^X Z^C K.$$

A log-log projection is called for by the equation

$$Y = AX^{CZ}$$

which turns first into

$$\log\ Y = (\log A) + CZ\ (\log X)$$

and then into

$$\log\ (\log Y - \log A) = \log C + \log Z + \text{log-log}\ X.$$

The projection of powers, roots, and reciprocals all find occasional use. Indeed any function whatever of two variables

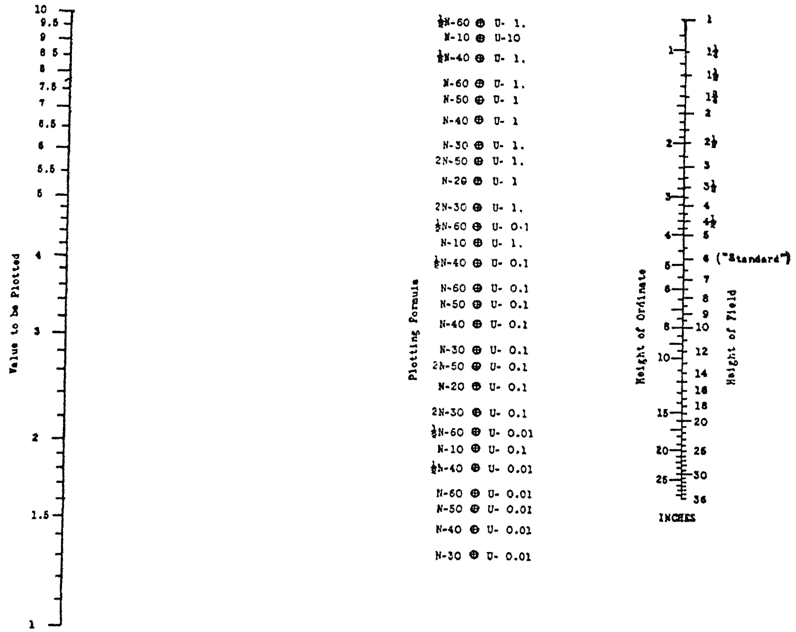
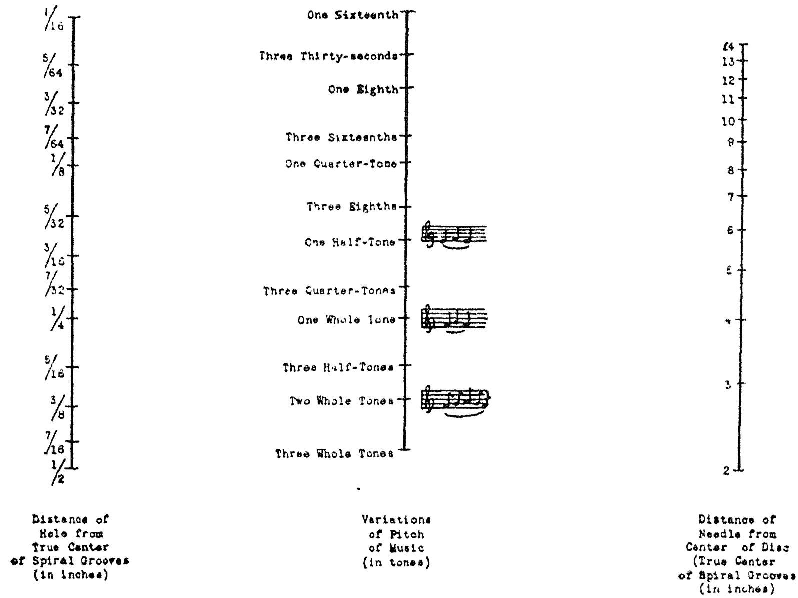


Fig. 431. Chart for Determining Scales of Curve-charts.

On the left-hand scale find the greatest value in the series and on the right-hand scale (inside) find the height at which it is to be plotted or (outside) the height of chart-paper. The nearest circle on the central scale, to a straight line between these two will give the side of the ruler (N) and the value of the ruler unit (U). Comp. with Fig. 407.



EFFECT OF OFF-CENTER HOLES IN PHONOGRAPH RECORDS ON PITCH OF MUSIC

Fig. 432. Parallel Nomograph Not Chartable by Formula.

The unknown variable, T (= change in musical pitch, in tones of the chromatic scale) cannot be plotted by the side of a slide-rule, or any other ruler. The formula of the chart is,

$$\frac{R + 2D}{R} = 2^{T/6}$$

in which R is the radius or distance of the needle from the true center of the disc (in inches) and D is the displacement of the hole therefrom (in inches). The formula can be stated as

$$\frac{2D}{R} = 2^{T/6} - 1$$

or $\log (2D) - \log R = \log (2^{T/6} - 1)$

or $\log D - \log R = \log (2^{T/6} - 1) - .3010$

If we lay off logarithmic scales of D and R through the desired ranges and then compute each value of these for each value of T which it is desired to plot, we can plot same by isopleths through the computed values of D and R. Thus:

T	$T\frac{\log 2}{6}$ or .05017 T	antilog $(T\frac{\log 2}{6})$ or $2^{T/6}$	$2^{T/6}-1$	R	$R(2^{T/6}-1)$ or 2D	D
set	.05017 a	antilog b	c − 1	set	d e	½ f
a	b	c	d	e	f	g
1	.05017	1.1225	1225	5.	.6125	.30625
2	10034	1 2599	.2599	5.	1.2995	.64975
½	.02508	1 05946	.05946	5.	2973	.14865
etc.	etc.	etc.	etc.	etc.	etc.	etc.

can be plotted if it can be reduced to the additive form $qx+pz=(p+q)y$.

The great difference between the nomograph and the curves described in the previous chapter, is that the curve of the latter has shrunk to a point in the nomograph, and the succession of curves has shrunk to a succession of points or single line. Incidentally the nomograph has shaken off the network of co-ordinates, though these are not essential even to the curve. In both of these steps the nomograph has reduced the labor of chart-making and increased the ease and accuracy of chart-reading. We have so far considered only the simple parallel nomograph, which is analogous to the simple parallel curves, but there are other nomographs which serve the purposes of the radiating curves and composite curve charts which we shall take up in the next chapter.

Chapter XLVII

ZIGZAG AND COMPOSITE NOMOGRAPHS

In a curious way the zigzag form of nomograph is even simpler than the parallel form. The parallel form has two outer axes and an inner axis which is slid along the base-line back and forth between the two outer axes as the scale moduli or coefficients of the variables on the outer axes are changed. In the zigzag form this central axis shrinks to a point—its own zero-point or origin on the base-line,—and having so dwindled moves back and forth along that base-line as a variable along a scale. It shrinks to a point because the third variable is turned into a constant. It lies upon the base-line because the new constant has been corrected for on an outer axis and one of them plotted reciprocally, that is, downward from the base-line. It moves back and forth along the base-line because the coefficient of the remaining independent variable has been turned into a new variable and hence has variable values. As a result, we must calibrate the base-line itself for the values of this new variable coefficient, or factorial variable, and lo and behold, we have a factorial chart without log projections, in many ways similar to the factorial radiating curve-chart for calculating formulae.

It is simplest, however, to explain the zigzag nomograph independently and from a different form of the geometrical theorem of similar triangles. We shall now speak of the base-line as an axis in itself, since it is calibrated, but it is to be understood that distances are not measured off upon it in units necessarily commensurable with the distances upon the other axes. But we anticipate.

If you lay off three lines or axes such that two are parallel and the third cuts through both, like the letter *N*, you can easily prove that along the three axes the distances cut off by a straight intersecting cross-line (in the finished chart, an isopleth) will have certain definite relations. In this case

we measure the distances from the intersections of the axes, that is, the two intersection points of the axes, are the origins of the axes. Let us call the first axis as before the x-axis and

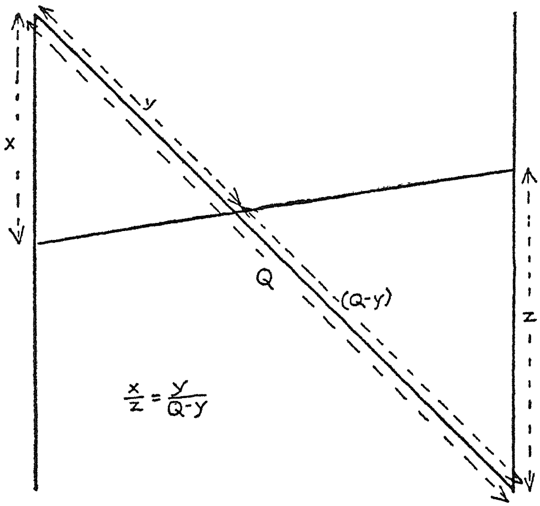


Fig. 433.

the distance laid off on it by the isopleth from the intersection of (x) and (y), as before, x. Let us call the middle or diagonal axis, the y-axis, measuring the distance, y, laid off on it by the isopleth, from the intersection of (x) and (y). Let us call the third axis, as before, the z-axis, measuring the distance, z, laid off by the isopleth, from the origin of the z-axis, that is, the intersection of the y and z axes. Now if we indicate the entire length of the y-axis, from x-origin to z-origin by Q to indicate that it is constant, we can quickly, from similar triangles, verify the following statement:

$$\frac{x}{z} = \frac{y}{Q-y}$$

Obviously if (with a chart-modulus, $m=1$ inch) we plot the values $x=X$, $z=Z$, and $\frac{y}{Q-y}=Y$, we may use this chart for calculations of the equation $Y=\frac{X}{Z}$, or $X=YZ$. As in the parallel nomograph, we note that space is conserved and accu-

racy gained by plotting the dependent variable upon the central or (y) axis.

The generalized equation is $$Y=\frac{AX\left|\begin{matrix}H_x\\L_x\end{matrix}\right|+B}{CZ\left|\begin{matrix}H_z\\L_z\end{matrix}\right|+D}+E.$$

As in the last chapter, we concern ourselves first with the two outer scales. Their ranges are:

$$R_x=H_x-L_x \qquad\qquad R_z=H_z-L_z$$

If we determine upon the height of the chart, Tm, or T inches, when the chart-modulus is 1 inch, then the scale-moduli are found from the two measures of the length of the scales

$$R_x m_x=Tm \qquad\qquad R_z m_z=Tm$$

$$m_x=\frac{Tm}{R_x} \qquad\qquad m_z=\frac{Tm}{R_z}$$

or, letting $m=1$,

$$m_x=\frac{T}{R_x} \qquad\qquad m_z=\frac{T}{R_z}$$

These are sufficient plotting instructions for the two outer scales if we are using a radiating scale-sheet.[1] But if we are using an engineer's rule, we shall want them turned into values of S (the number of scale-moduli per inch):

$$S_x=\frac{r_x}{t_x} \qquad\qquad S_z=\frac{r_z}{t_z}$$

and if S does not at once show a ruler-copying value, that is, 1, 2, 3, 4, 5, 6, or decimal multiple or submultiple thereof, we shall alter R and T a bit until it does for each scale.

Now we can at once lay off the central or y-axis. Care must be taken to direct it at the true origins of the outer axes, which will differ from the apparent origins by the amount of $\frac{B}{A}$ (on the x-axis) and $\frac{D}{C}$ (on the z-axis). The entire chart should be about square in outline (unlike the parallel nomograph) or even slightly narrower, to get the best results in

[1] That is, a sheet facilitating scale adjustment by triangulation. These sheets are found in Professor Lipka's book and are described in the previous chapter.

reading from its isopleths. The scale for Y can be inserted by solving the equation for different values of the variables and plotting them on the y-axis by isopleths. The scale is not uniform, it is a variety of reciprocal projection, each point having a value proportional to the ratio of the segments of the line on either side of the point. Every value calibrated on the scale must therefore be individually computed and plotted by isopleths. This is a simple way to construct a zigzag nomograph.

Since, however, the projections through these points or values from any point, n, on the z-scale, forms as we know, an arithmetical scale on the x-axis, it is a simple matter to reverse the process and plot a temporary working scale (which we may call W) arithmetically along this x-axis, calibrated equal to Y and such that from it we may easily project the Y-scale on the y-axis.[2] We select any convenient point, n, preferably near the middle, on the z-axis. Through two known points already computed and plotted on the central scale, we project isopleths from n to the x-axis. Then we lay off a complete scale about these two points and with n as center, plot their projections on the y-scale; lastly we erase the temporary scale and the point n. This is a better way to construct a zigzag nomograph.

The student will seek a mathematical expression for the plotting of this eccentric y-scale. Now to find the values of y in the equation $\frac{x}{z}=\frac{y}{Q-y}$, we must first examine the true values of x and z. These are the distances along the two axes from the true origins, which differ from the apparent origins by the amounts of $\frac{B}{A}$ and $\frac{D}{C}$. So to be quite correct we must write

$$x=m_x(X+\frac{B}{A}) \qquad\qquad z=m_z(Z+\frac{D}{C})$$

[2] The projections of Y upon the x-axis from any point in the z-axis are always regular (i.e., uniform or arithmetical, or in accordance with the scale of X), because X itself is regularly laid off, and by the formula, Y varies directly with X when Z is taken constant. The temporary working-scale, W, cannot be plotted upon the z-axis from a point on the y-axis, because, by the formula, Y varies inversely with Z when X is taken constant; hence, the transversals from uniform intervals along the z-axis would only project upon the y-axis a sort of reciprocal scale projection thereof.

This is true from the definitions of m_x and m_z as scale moduli. So

$$x = \frac{m_x}{A}(AX+B) \qquad\qquad z = \frac{m_z}{C}(CZ+D)$$

Substituting these in the equation $\frac{x}{z} = \frac{y}{Q-y}$, we have

$$\frac{y}{Q-y} = \frac{Cm_x(AX+B)}{Am_z(CZ+D)}$$

Now $\frac{AX+B}{CZ+D} = Y - E$, so we may write

$$\frac{y}{Q-y} = \frac{Cm_x}{Am_z}(Y-E) = \frac{Cm_x(Y-E)}{Am_z}$$

$$\frac{Q-y}{y} = \frac{Q}{y} - 1 = \frac{Am_z}{Cm_x(Y-E)}$$

$$\frac{Q}{y} = \frac{Cm_x(Y-E)+Am_z}{Cm_x(Y-E)}$$

$$y = \frac{Cm_x(Y-E)}{Cm_x(Y-E)+Am_z}Q$$

This is a cumbersome expression and shows that it is simpler to compute the y-scale empirically as before described. If we write

$$y = m_y(Y-E)$$

then we see that m_y, the modulus for the y-scale is

$$m_y(Y-E) = \frac{Cm_x(Y-E)\,Q}{Cm_x(Y-E)+Am_z}$$

$$m_y = \frac{Cm_x\,Q}{Cm_x(Y-E)+Am_z}$$

from which we see that this modulus is not constant but changes with the values of Y, the variable. This is merely algebraic proof of the irregular nature of the y-scale projection. We may observe from the equation for y that it is a fraction

of the constant distance, Q, between the true origins of the outer scales and that y (and hence the scale for the Y variable) will always lie between these origins (and hence between the two outer scales) so long as the denominator of the fraction exceeds the numerator; we likewise observe that the y-scale will lie outside of the two parallel scales when the numerator exceeds the denominator.

An interesting thing about the y-scale is its Y-value or calibration at the point midway between the two outer scales, that is, when $y = \frac{1}{2} Q$. For this we write

$$y = \tfrac{1}{2} Q = \frac{Cm_x(Y - E)\, Q}{Cm_x(Y - E) + Am_z}$$

$$Cm_x(Y - E) + Am_z = 2Cm_x(Y - E)$$

$$Am_z = Cm_x(Y - E)$$

$$Y - E = \frac{Am_z}{Cm_x}$$

$$Y = \frac{Am_z}{Cm_x} + E = \frac{A/C}{m_x/m_z} + E$$

$$= \frac{A/m_x}{C/m_z} + E$$

Thus at the point midway between the two scales the value of Y is always easily found, by dividing the constant coefficient of each independent variable by its scale modulus, then dividing the quotient for x by that for z and adding any added constant. If we have the simple case in which both independent scales were plotted on the same moduli, and there is no added constant, then the midpoint of the (y) scale expresses the ratio of the coefficients in the formula. If the coefficients are alike but the moduli are different, and there is no added constant then it expresses the ratio of the moduli. Of course the addition of a constant, E, merely raises these values by its amount. And when coefficients and moduli are alike, and no constant is added, the midpoint has the value of 1, and at equal distances on either side all the other calibrations will be found to be mutually reciprocal.

Useless as is the mathematical expression for y and m_y, a similar expression for the temporary projection of the (y)

scale upon the x-axis, by means of which the y-scale can be plotted without computing, is valuable. We select on the z-scale (preferably near its center) a fixed point, n, the calibration or Z-value of which let us call N. If we run transversals through N and every uniform value of Y to be calibrated on the y-scale, we know that the transversals would mark off a regular scale upon the x-axis. Let us call the temporary working scale w. The intervals or scale modulus, m_w, of the w-scale, would have the same relation to the modulus of x as the calibrations of X have to the calibrations of W. Thus $\frac{m_w}{m_x}=\frac{X}{W}=\frac{X}{Y}$. Now when $Z=N$, the value of $\frac{X}{Y}$ is as follows.[3]

$$AX+B = (CN+D)\ (Y-E)$$

$$X=\frac{CN+D}{A}Y-\frac{CN+D}{A}E-\frac{B}{A}$$

Drop the added constants since they merely shift the zero point and

$$X=\frac{CN+D}{A}Y$$

$$\frac{X}{Y}=\frac{CN+D}{A}=\frac{X}{W}$$

$$\frac{m_w}{m_x}=\frac{CN+D}{A}$$

$$m_w=\frac{CN+D}{A}m_x$$

Or if we wish to work with the length, n, of the plotted point, from the true z-axis origin, instead of its calibrated Z-value, N, we have

$$n=m_z\ (N+\frac{D}{C})=\frac{m_z}{C}(CN+D)$$

$$m_w=\frac{C}{A}\frac{m_x}{m_z}n.$$

If we wish the modulus in terms of the ruler to use, i.e., the number of moduli per inch (or per chart-modulus, whatever it be), we have

$$S_w=\frac{1}{m_w}=\frac{AS_x}{CN+D}=\frac{AS_x}{CS_z}n.$$

[3] The mathematical steps here are outlined without full details, as the latter would make the equations more cumbersome than their importance justifies.

In short, to prepare a zigzag nomograph for the equation

$$Y=\frac{AX\left|\frac{H_x}{L_x}\right|+B}{CZ\left|\frac{H_z}{L_z}\right|+D}+E,$$

we need only compute the following expressions in order to plot with an engineer's hexagonal rule:

$S_x=\frac{r_x}{t_x}$ in which r_x approximates $R_x=H_x-L_x$

$S_z=\frac{r_z}{t_z}$ in which r_z approximates $R_z=H_z-L_z$

$S_w=\frac{AS_x}{CS_z}n$ in which n is the distance (in inches or units of the chart modulus) of a fixed point on the z-axis from the origin thereof,

or

$S_w=\frac{AS_x}{CN+D}$ in which N is calibrated Z-value on the z-scale.

The usefulness of the above expressions is in the search for ruler-copying values of S which will enable us to plot directly from the engineer's rule. For this purpose the same tabular arrangement of columns for the values of t_x, r_x, S_x, AS_x, n, S_w, CS_z, S_z, r_z, and t_z, should be made in order that slightly different values of t and r may be tried on various scales. If, however, we work with a radiating scale-sheet, then the scale-moduli are wanted and these are as follows:

$$m_x=\frac{T}{R_x}$$

$$m_z=\frac{T}{R_z}$$

$$m_w=\frac{Cm_x}{Am_z}n$$

or

$$=\frac{CN+D}{A}m_x$$

The chart will also express the exponential equation

$$BX^A=(EY)^{(CZ+D)} \text{ or } \frac{A\log X+\log B}{CZ+D}=\log Y+\log E$$

and amplifications thereof; here one outer scale is logarithmically, the other arithmetically projected. When the other

functions of the variables are used, such as powers, roots, or reciprocals, the corresponding projections may be called for. In all this work the chart will be seen to handle added constants with much less trouble than the factorial or logarithmic

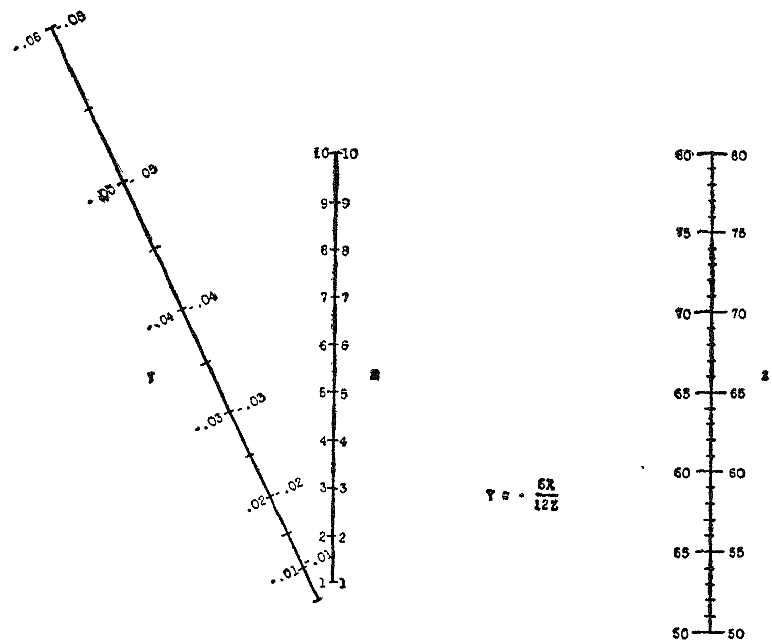


Fig. 434. The Y-Scale Outside the Parallel Scales.

Plotted with $S_x = -2$ (the negative sign shows that scale-values increase upward along this axis, instead of downward as normally), $S_z = 6$, and $S_w = .01$ for $N = 33\frac{1}{3}$. The positive sign of S_w shows that W-values (on the X-axis) increase (that is, become larger positive or smaller negative values) downward, as normally; this consequently applies also to the Y-values on the Y-axis. The upward or reversed projection of the X-scale (due to the negative sign of S_x) shows that the Y-axis, passing through the true zeros (plotting origins) of the parallel axes, lies below and outside the two scales, hence the Y-scale is outside the parallel scales.

This form is not of much value. The dependent variable would better be X when it is used.

(NOTE: Above dimensions as of original drawing, here reduced to half-size.)

parallel nomograph, for it does not require a specially computed scale for the shifted zeros. The zigzag nomograph is, however, on the whole of less value than the parallel nomograph, for the central axis, being diagonal, is often crossed by the isopleths at very small angles, and the readings naturally become less accurate.

Many other forms of nomographs have been devised beside the parallel and zigzag forms. The theory and making of these involve mathematical work and detail outside the scope of this book. They are based upon various geometrical theorems

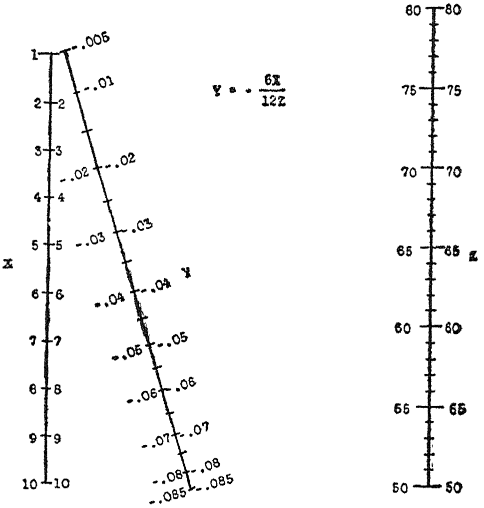


Fig. 435. The *Y*-Scale Inside the Parallel Scales.

The same equation as in Fig. 434, here plotted with $S_x=2$, $S_z=6$, and $S_w = -0.01$ for $N=33\frac{1}{3}$. The positive signs of S_x and S_z show that the normal directions of plotting of the X and Z scales obtain, X increasing downward and Z upward and the Y axis consequently passing between them. The negative sign of S_w shows that the W-values along the X-axis and therefore the Y-values along Y-axis, increase upward instead of downward (that is, grow larger positively or smaller negatively).

This form is usually better than that in Fig. 434, for though it gives less detail to parts of the Y-scale, it places the dependent variable inside, giving more accurate readings.

(Note: Reduction of half-size.)

and are generally built up with straight lines for scales on the sides of imaginary triangles and parallelograms. It is indeed possible to have nomographs with curved axes but these are not often encountered, nor is their need more than exceptional. A very large body of the less used nomographs are proportional, and can be used for equations containing four variables which are in or can be put into the form of a proportion. These charts use two isopleths either parallel, perpendicular, or with intersections upon a dummy line, in order to afford the readings for the four variables. The interest attaching to these less common types of nomographs is still largely

academic; the two simple forms, the parallel and zigzag, afford adequate calculating facilities for all practical purposes.

We have so far considered only the simple forms of these nomographs, in which there appear but a single set of three axes, and which are suitable only for equations with three variables (two unknown). The most interesting form of nomograph is a compound one composed of two or more interlocking single nomographs and suitable for equations with more than two independent variables. Each single nomograph is a set of three axes, but when two or more sets are combined, one of the axes of each set serves double duty,

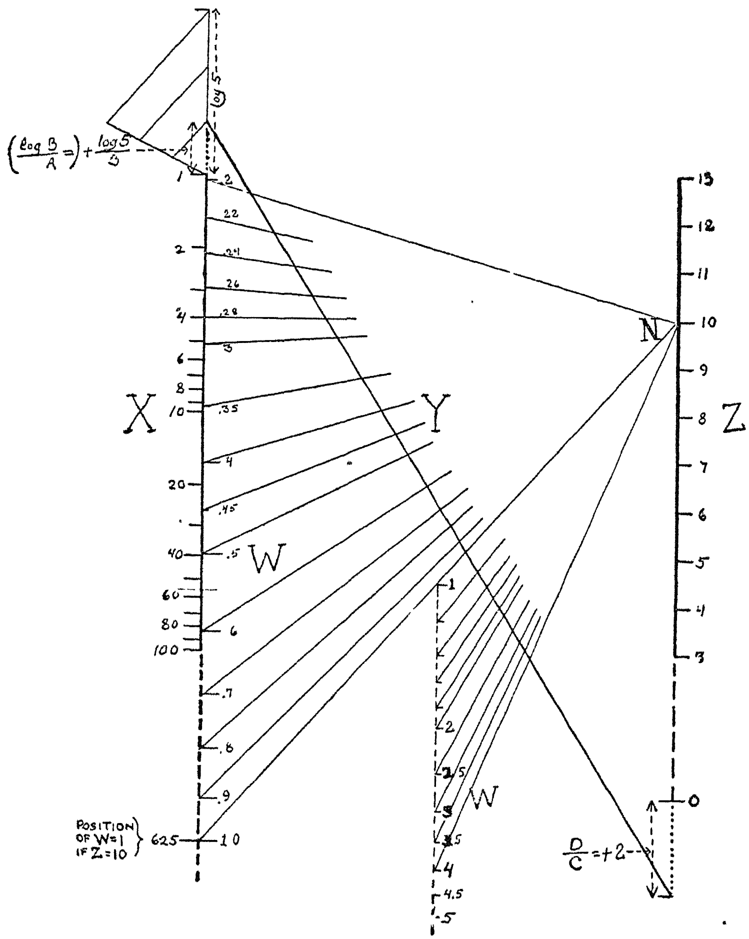


Fig. 436. Construction of Factorial Zigzag Nomograph—Unfinished.

NOTE TO FIG. 436

Showing the working-scale, W, from which the dependent variable scale, Y, is plotted; and the true origins of the independent variable scales, shifted for added constants.

The equation is $$5a^3 = (5.7171b)^{c+2}$$

in which a varies from 1 to 100, c varies from 3 to 13, and b is to be found. Let $X = a$, $Z = c$, and $Y = b$.

Since $$\log 5 + 3 \log a = (c + 2)(\log 5.7171 + \log b)$$

or $$\frac{3 \log a + \log 5}{c + 2} = \log b + \log 5.7171$$

and the typical formula is

$$\frac{A \log X \left| \begin{array}{c} Hx \\ \\ Lx \end{array} \right| + \log B}{C Z \left| \begin{array}{c} Hz \\ \\ Lz \end{array} \right| + D} = \log Y - \log E$$

we have, by substitution

$$\frac{3 \log X \left| \begin{array}{c} \log 100 \\ \\ \log 1 \end{array} \right| + \log 5}{Z \left| \begin{array}{c} 13 \\ \\ 3 \end{array} \right| + 2} = \log Y - (-\log 5.7171)$$

If we wish the chart to have a total height of $T = 5$ inches we can plot with the following scale moduli:

$$Rx = Hx - Lx = \log 100 - \log 1 = 2 - 0 = 2 \qquad Rz = Hz - Lz = 13 - 3 = 10$$

$$mx = \frac{T}{Rx} = \frac{5}{2} = 2.5 \text{ inches} \qquad mz = \frac{T}{Rz} = \frac{5}{10} = .5 \text{ inches}$$

We must plot the x-scale from a logarithmic scale having one deck for every $2\frac{1}{2}$ inches, and the z-scale from the 20-side of an engineer's ruler, after allowing for the added constant in each case.

If we select as the fixed point, N, for our working scale, W, the point calibrated as 10 on the z-scale

$$m_w = \frac{C N + D}{A} mx = \frac{10 + 2}{3} (2.5) = 10 \text{ inches}$$

and we plot the w-scale from an inch-rule (the 10-side of the engineer's ruler).

To position the W-scale we calculate the value of X for any value of Y we choose: thus,

$$\left. \begin{array}{r} \text{when } Z = N = 10 \\ \text{and } Y (= W) = 1 \end{array} \right\} \log X = \frac{(Z + 2)(\log Y + \log 5.7171) - \log 5}{3}$$

$$= \frac{12 (\log 1 + .75722) - .69897}{3}$$

$$= 2.7959$$

$$X = 625.0$$

While this value of X lies outside the range, and is therefore inconvenient, we need not recompute X for another value of Y, but merely extend the X scale sufficiently to plot $W = 1$, after which other W values follow by the ruler.

being common to two sets and effecting the combination. Thus if we have the formula or equation $A = B+C+D+E$, we will have to break the right side of the equation, having

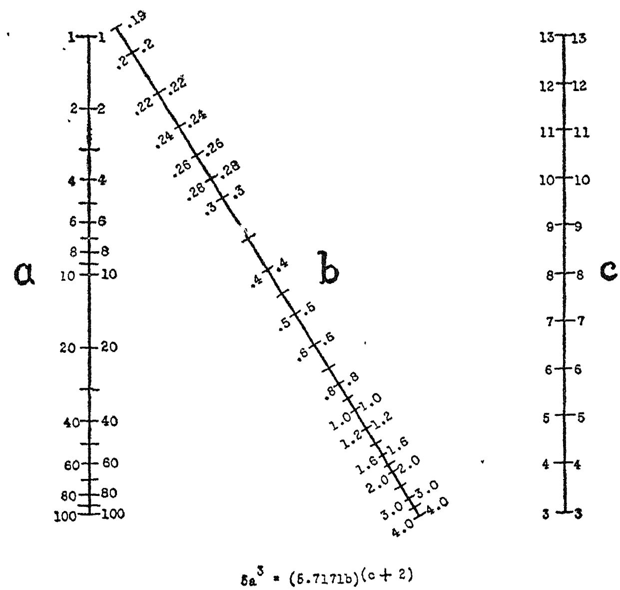


Fig. 437. Construction of Factorial Zigzag Nomograph—Finished.
The temporary working scale, W, is erased after the calibration of the y-scale therefrom. Also the extensions of scales to the true origins and to $x = 625$ have been erased. Scales have been calibrated on both sides to facilitate readings when using an opaque ruler as isopleth.

four independent variables, into two groups of two each and make a parallel additive nomograph of each group, adding a third axis to each to express the resultant of each group, and then we can combine the resultants in a third parallel nomograph to show the dependent variable, A. We would write

$$f = B+C$$
$$\text{and} \quad g = D+E$$
$$\text{and} \quad A = f+g$$

In the first two groups we might let f and g be middle axes, but in the third group we would use them as outer axes writing A as the middle axis of the group. The order of the axes would be B, f, C, A, D, g, E. If for convenience we wished A to be the final axis, then we should have to fall back on the use of inverted x-scales and carefully arrange the scales so that

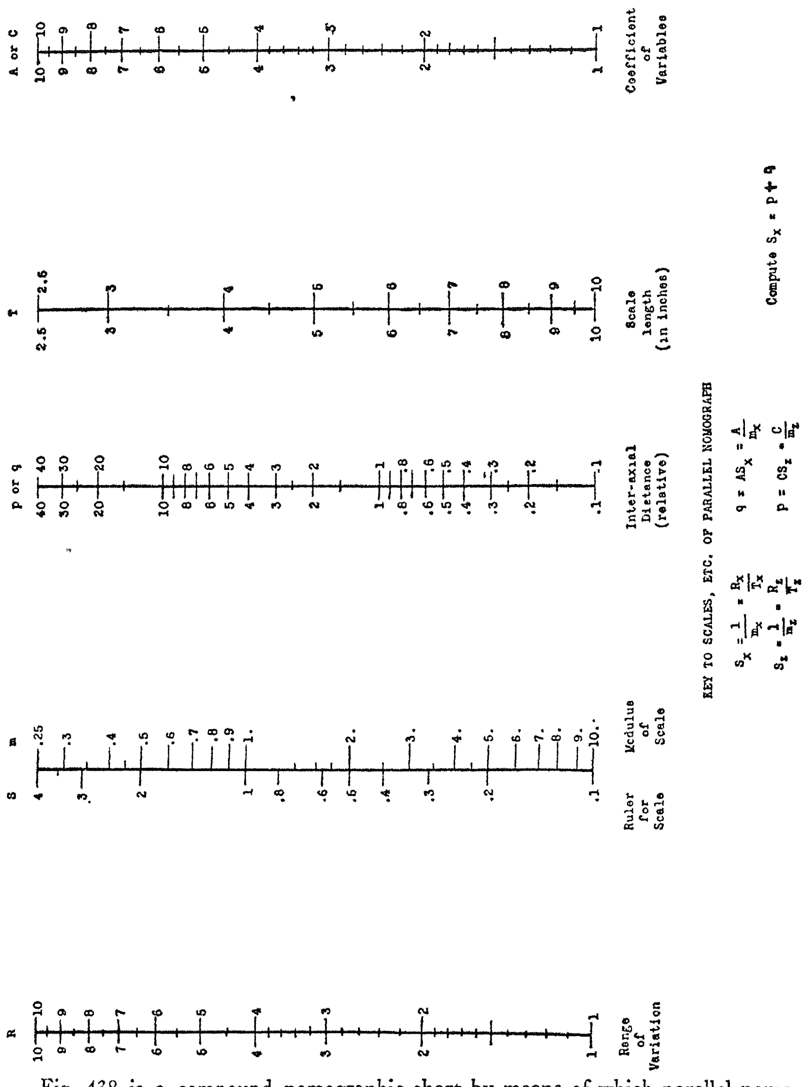


Fig. 438. Chart to Construct Parallel Nomographs.

Fig. 438 is a compound nomographic chart by means of which parallel nomographs may be constructed.

Draw isopleths from Rx and Rz on the R scale (first) to Tx and Tz on the T scale (fourth) and find scale moduli, mx and mz and rulers Sx and Sz on the Sm scale (second). From the latter draw isopleths to A and C on the fifth scale and read the values of p and q on the third scale. Add the latter, p and q, to get Sy. In the above Rx and Rz are the ranges of the two independent variables, x and z; Tx and Tz are the tenths (in inches) to be given these scales on the chart; Sx and Sz are the engineer's rules (or number of units per inch) to use in plotting them, A and C are the coefficients of the two variables, x and y; and p and q are the horizontal distances between axes (p between x and y, q between y and z). Sy is the engineer's rule (or number of units per inch) to use in plotting the dependent y-scale. Position the scales (for added constants) by a single trial isopleths.

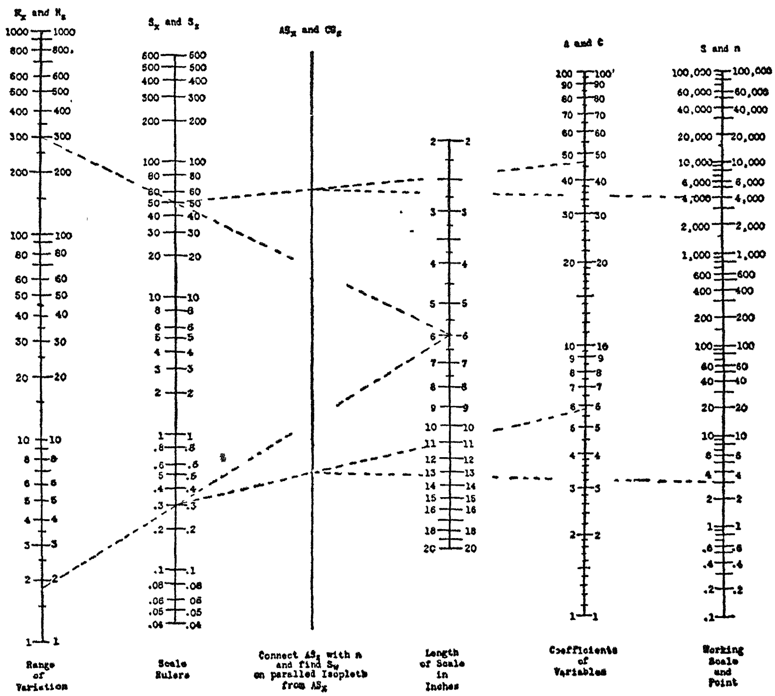


Fig. 439. Chart to Construct Zigzag Nomographs.

Here is a compound nomographic chart by means of which zigzag nomographs may be constructed. Draw isopleths from R_x and R_z on the first scale to T_x and T_z on the third scale, and read S_x and S_z on the second scale. From the latter draw isopleths to A and C on the fourth scale and note intersected points on the dummy axis. From the last, the intersection of the dummy scale and the isopleth through S_z and C, draw an isopleth to n on the fifth scale. From the other dummy axis intersection, $A\,S_x$, draw a parallel isopleth to the fifth scale and read S_w. In the above A and C are the coefficients of the independent variables, x and z; R_x and R_z are their ranges; T_x and T_z their scale lengths in inches; n is the fixed point distance on the z-scale to project the working scale, W, on to the y-axis as Y, and S_x, S_z, and S_w are the sides of the engineer's rule (or number of units per inch) to use in plotting. Position the scales for added constants by means of these plotting-units.

while each y and z scale went in the same direction, each x scale went in the opposite one. Compound nomographs can be used for equations with many factors instead of terms, in precisely the same way, merely using logarithmic projection or zigzag nomographic form. The sub-total axes (f and g in the example just cited) or the sub-product axes in factorial nomographs, are generally left without calibrations, as no one is interested in reading their values. They are necessary merely as fixation points secured by the first interpolation and fixing the isopleth for the next step. They are called dummy axes.

The fact that the zigzag nomograph performs multiplication and addition on arithmetically projected scales makes it useful for compound nomographs of formulae involving both addition and multiplication. This, indeed, is the chief reason for the importance of the zigzag form. Thus an equation of the general type $A = BC + DE$ can be solved by the use of two zigzags for the two multiplication processes and a parallel for the sum of their products. This equation could not be shown on parallel nomographs alone, because in them logarithmic projections would have been necessary for the factorial processes and the addition of the products, were logarithmically projected would have shown not a sum but a third product.

It has already been said that many other projections can be used beside logarithmic and arithmetic ones. Squares, cubes, roots, and trigonometric functions can be used. When such functions are used, the equations $px = mX$ or $x = m_x X$ no longer hold, but must be modified to $px = m\,f(X)$ and $x = m_x f(X)$. This will require the modification of the calculating formulae which have been given for the scale-moduli, but the procedure is so similar that it may be left to the devices and ingenuity of the chart-maker. Nomograph-making presupposes a fairly thorough understanding of the equation to be plotted and with this as a basis, the ingenious experimenter will find various and adequate methods of charting.

Upon the finished chart the scales should be provided with titles below or above them, explicitly stating the variables to be located or read on each scale. The formula which the chart expresses should also be available to the reader somewhere about the chart. The best mechanism for the reading of the scales is a strip of transparent celluloid with a fine straight

line drawn in ink upon its lower surface. A straight-edge or ruler, if possible with a transparent edge, can be used; and in an emergency a piece of thread can be drawn tight and held for the readings.

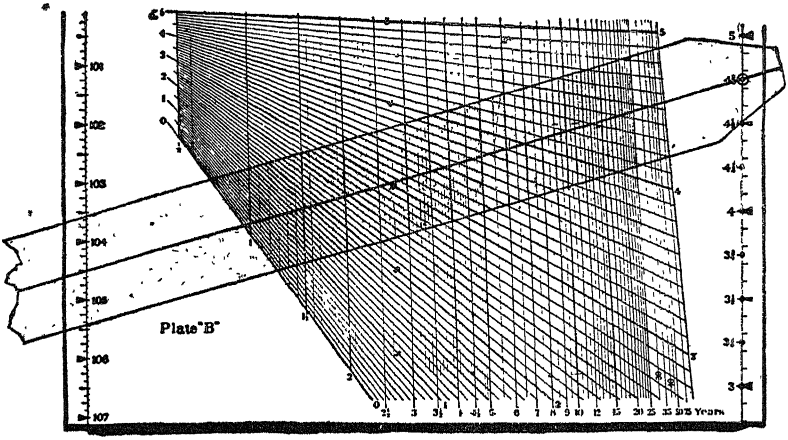


Fig. 440. In Quadratic and Cubic Equations the Position of the Central Axes Becomes Variable, and a Chart-field Takes the Place of a Single Scale.

This is the Darville-Johnson Bond-Yield Chart, a chart for determining quickly the yields of all types of bonds, including premium bonds, maturing in any number of years at practically all coupon rates now in use, including odd fractions. —*Published by Prentice-Hall, Inc.*

Whole books have been written about the nomograph and while it is still a little known chart, yet it is fast increasing in popularity, and deservedly so. It is the most easily constructed and accurately read of all calculating charts, and the results which can be accomplished with it are always a source of amazement to the uninitiated.

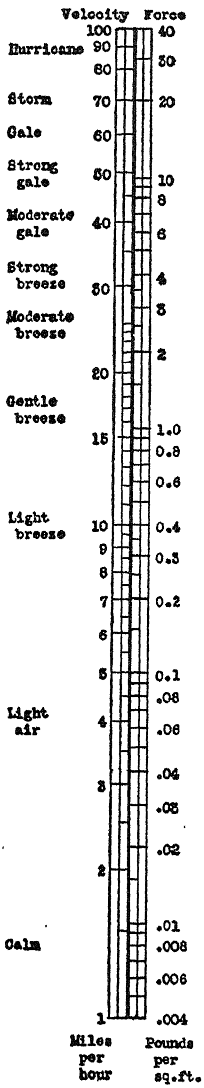
CHAPTER XLVIII

SLIDE RULES

A calculating device which is even more simple to operate than the alignment chart or nomograph, but may be considered closely related to it, is the slide-rule. We have seen that by the use of special projections, the curve of an equation may often be straightened, increasing both the ease and accuracy of calculations based upon the curve. We have seen that in the nomograph the chart field has also been eliminated with still further simplicity and benefit. But we come now to the slide-rule, in which even the straight-edge (that rudimentary substitute for the curve) has been eliminated, the scales losing their fixed position with regard to each other and being freely movable. The slide-rule is therefore nothing more than two movable scales along the same axis, that is, in contact with each other.

We can, however, approach the subject of slide-rules even more simply if we consider first the stationary rule. The stationary rule is nothing more than a single axis bearing two or even more scales, one upon each side of the axis. This is the graphic chart of equations containing but two variables, (only one independent or known variable). The scales are so adjusted that for every value of one variable, calibrated on one scale the corresponding value of the other variable may be read upon the other scale at precisely the same point along the axis. Fixed or stationary scales may be laid off upon logarithmic, arithmetical or any other projections or combinations of projections. Such charts are useful in the place of small conversion tables, but must be made large or in several segments when detailed readings are necessary. They can be used as ready reckoners for foreign exchange, temperature equivalents in Fahrenheit and Centigrade, the conversion of metric and common systems of measures and weights, and an infinite

variety of similar cases in which the relation between two variables is constant and fixed.



THE FORCE OF WINDS (Standard Table)

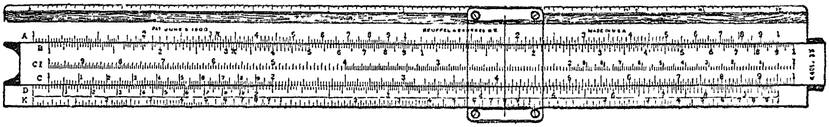
Fig. 441. A Stationary or Fixed Rule.

As its name suggests, the slide-rule is not fixed or stationary. If you will take two ordinary rulers, one with its scale upon the upper edge, and the other with its scale upon its lower edge, and bring them together so that the two calibrated edges will fit together, you will have the simplest form of slide-rule. Calculating is done by the simple trick of sliding one ruler along the other and reading the corresponding values in the new positions. Thus in order to add 2 and 4, you need only slide the upper rule along the lower one, until the zero point on the upper rule is over the figure 4 on the lower rule, and then read the figure on the lower rule below the figure 2 on the upper rule. Obviously you have in this way added two inches to the four inches on the lower rule and you will get six inches on the lower rule. The upper rule merely tells you how much you have added to the original distance on the lower rule. Likewise to subtract 2 from 6 you need merely place the 2 on the upper rule over the 6 on the lower rule and read back to the figure 4 on the lower rule under the 0 on the upper rule. This amounts to deducting 2 inches from the original 6 inches on the lower rule, giving you a remainder of 4 inches on the lower rule. In short, the slide-rule is merely a device for the direct addition or subtraction of distances.

In the device just explained, the calibration of the two rulers forms an arithmetical series and hence the calculating power of this device is limited to the processes of addition and subtraction. In order to use the device for the processes of multiplication and division, of course it is only necessary to calibrate the rules upon logarithmic projections, so that the addition or subtraction of

the logarithmic distances will indicate the processes of multiplication and division of the numbers appearing on the scale. As in the case of the nomograph, the calibration can be an arithmetic or logarithmic projection of sine, tangent, square, cube, root, and other functions of numbers as well as of the numbers themselves.

The ordinary commercial slide-rule is nothing more than a series of these scales of logarithmic projections of various functions, mounted upon bits of wood which fit closely together and can be conveniently handled. One rule is made

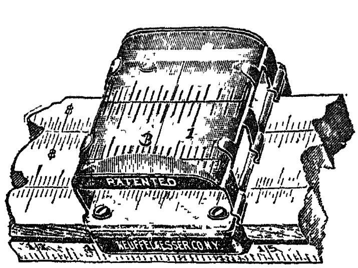


Courtesy of Keuffel & Esser, N. Y.

Fig. 442. A Slide Rule.

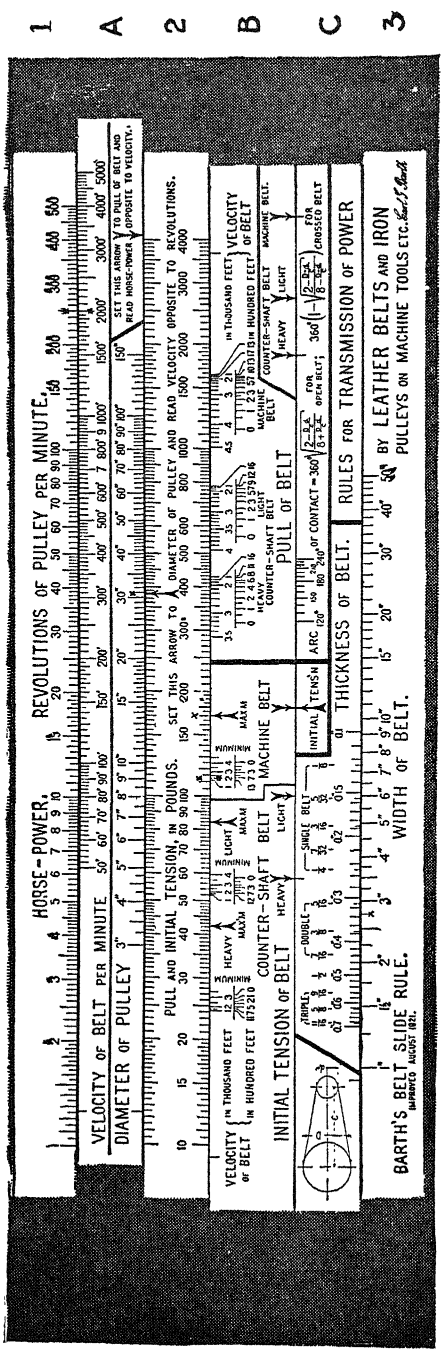
For multiplication of numbers, squares, cubes, tangents, sines and other circular functions and also showing logarithms of numbers.

much smaller and fits within a groove on the other, sliding freely back and forth along that groove so that it is not necessary to hold the two rules constantly together. They are so tightly adjusted that the two rules remain without shifting in whatever positions they are placed, leaving you free to take the readings on the scales with great care. The inner rule is called the "slide." A "runner" is also attached to the outer rule for convenience in taking readings, being generally a small piece of glass on which a fine hair-line has been drawn at right angles to the scale. When you have positioned this runner so that the hair-line crosses the point desired upon one scale, the hair-line will also cross the desired point upon the other scale, and the reading on the second scale can be more



Permission of Keuffel & Esser, N. Y.

Fig. 443. The Magnifiers Increase the Accuracy of Readings.



From "The Improved Belt Slide Rule," by Carl G. Barth, in Management Engineering, by permission.

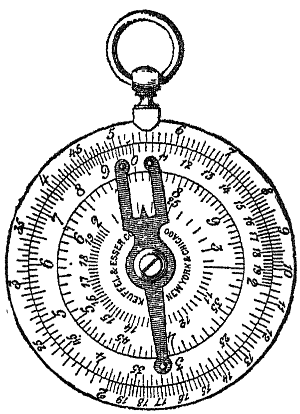
Fig. 444. Slide Rule With Three Slides.

Designed by Mr. Carl Barth for Belting Measurements.

The slides are marked A, B, and C.

easily taken. And when the two scales which you are using are not in immediate contact, but are parallel some distance from each other, the runner is necessary to project the desired point from the first scale to the second, forming a sort of ordinate across the two scales. Magnifying glasses are often attached to these runners so as to facilitate more exact readings.

Because the construction of a straight slide-rule calls for rather delicate carpentry,[1] slide-rules on short notice in the home or office are more easily made in circular form. And because a circle is endless and over three times as long as its diameter, the circular slide-rule can be made on much larger scale and with consequently greater accuracy than a straight slide-rule of the same physical size. For the circular slide-rule, you need merely pin together through their centers two



Courtesy of Keuffel & Esser, N. Y.

Fig. 445. A Circular Slide Rule—Pocket Size.

Owing to the great length of a circumference (compared to a diameter), and to the overlapping (because endless) edges, the circular rule is very compact.

circular pieces of paper, the smaller one uppermost, so that a scale can be drawn on the visible inner edge of the lower one which will always be in contact with a scale drawn on the outer edge of the upper disc. Then by rotating one disc above the other the readings can be taken off in precisely the same way as with a straight slide-rule in which one rule was slid along the other. The circular slide-rule operates on precisely the same principles as the straight slide-rule, adding or sub-

[1] Excellent examples of straight slide-rules for special purposes may be found in the writings of Mr. Carl Barth. The circular slide-rule has been more used by Mr. Walter N. Polakov.

tracting distances around a circumference instead of along a straight line, the distances on the scales being prepared by angular instead of linear measurement.

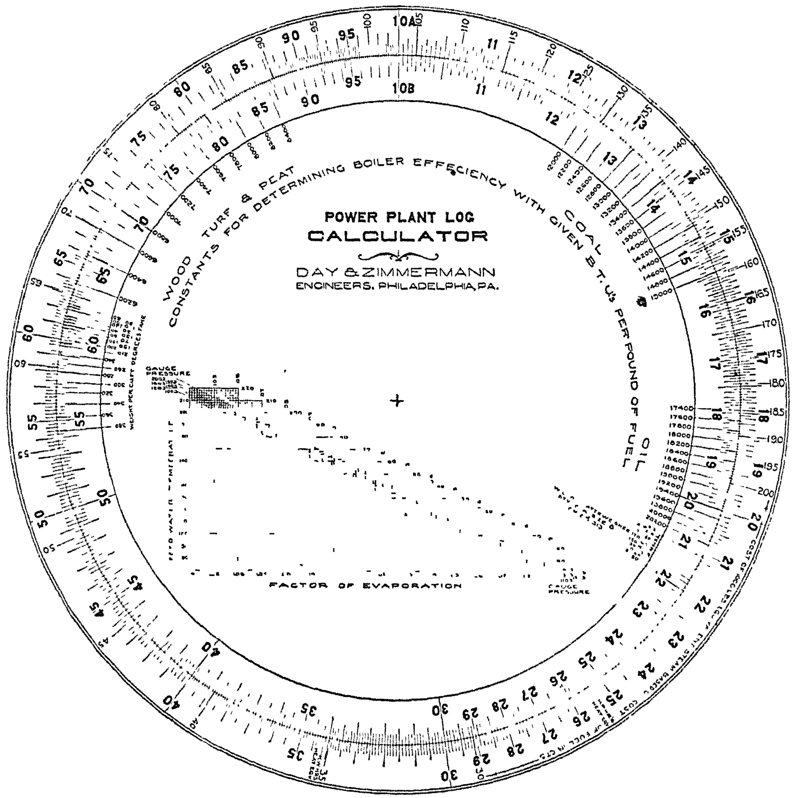


Fig. 446. A Special Circular Slide-Rule.
Devised by Mr. Walter N. Polakov.—*Permission of Mr. Polakov.*

Some difficulty may be met in the calibration of the circular scale. It is comparatively easy to project any scale or calibration upon a straight line, but to project it upon a circular line or upon an arc of a circle, it is necessary to use an instrument for measuring angles, called a protractor. Protractors are almost invariably calibrated in degrees, the entire circle being divided into 360 degrees. Two other units of angular measurement are known, grades (6400 grades to the circle) and radians (the radian being the arc equal to the radius of the circle), but neither of these is any more useful for

the purpose in hand than the degree. Even the metric system has no decimal unit of circular measurement. This is unfortunate because a decimal circular measurement system would often be convenient. Sometimes circles are divided into one hundred parts for the plotting of 100% circles or pie-charts but these are not usually of sufficient accuracy and precision to use as protractors. Your best plan in the making of circular scales is to decide beforehand approximately how far around the circle you wish your scale to run and then turn your scale-distances which you would use in calibrating the straight line scale into the nearest convenient number of degrees and lay them off with a large protractor or scale of degrees, last of all re-calibrating the scale for the desired value from your conversion table. Thus, if you wish a scale which runs from 0 to 10 in actual distances to extend about a quarter of the way around the circle, you can plot your table of distances directly onto the circle from your protractor by using that portion of the protractor which extends from 0 degrees to 100 degrees, but if you wish your scale to extend over half way around the circle you must first double the actual distance values before plotting them as degrees, so that you can plot through the protractor from 0 degrees to 200 degrees.

Circular slide-rules can be made with a number of independent scales, each sliding on separate pieces of paper but all pivotted together at their centres by a small rivet. A substitute for the runner can be attached in the form of a strip of transparent celluloid with a fine ink line drawn radially from the center or pivotal point. This ray can then be swung about the circle and laid over any desired point on the scale to facilitate readings in the same way as the runners on a straight slide-rule. Another device is to make the uppermost circular sheet of paper so large as to cover all the other sheets and then cut windows or circular slits in the upper sheet where the lower scale should be seen and mark small pointers or arrowheads next to the windows for readings on the lower scale.

Circular slide-rules are easily made when you have once grasped the fundamentals of their construction and they afford the greatest play for ingenuity. When skilfully made, they perform the most intricate calculations with astonishing ease and simplicity. Their only limitation appears to be that they perform in one operation only that particular type of mathe-

matical process for which they have been designed. As they operate by adding and subtracting distances, they will perform

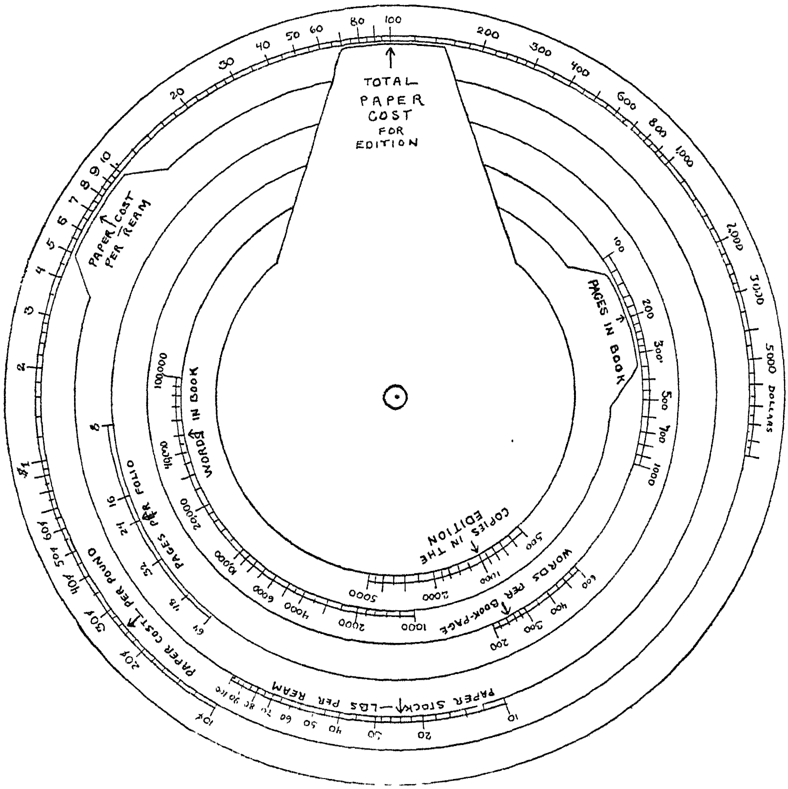


Fig. 447. A Circular Slide Rule with Many Variables.
Showing Cost of Book Printing.

addition and subtraction if the calibration is in terms of arithmetical series, that is, in the natural numbers. They will perform multiplication and division of as many factors as there are scales, by the simple trick of calibrating them logarithmically, that is for the logarithms of the natural numbers. But they cannot be used at the same time for addition and multiplication, for the two processes require different types of calibration or projection of scales. This limitation is ordinarily not a serious one, because most tedious business compu-

tations are either processes of addition or multiplication but not combinations of the two.[2]

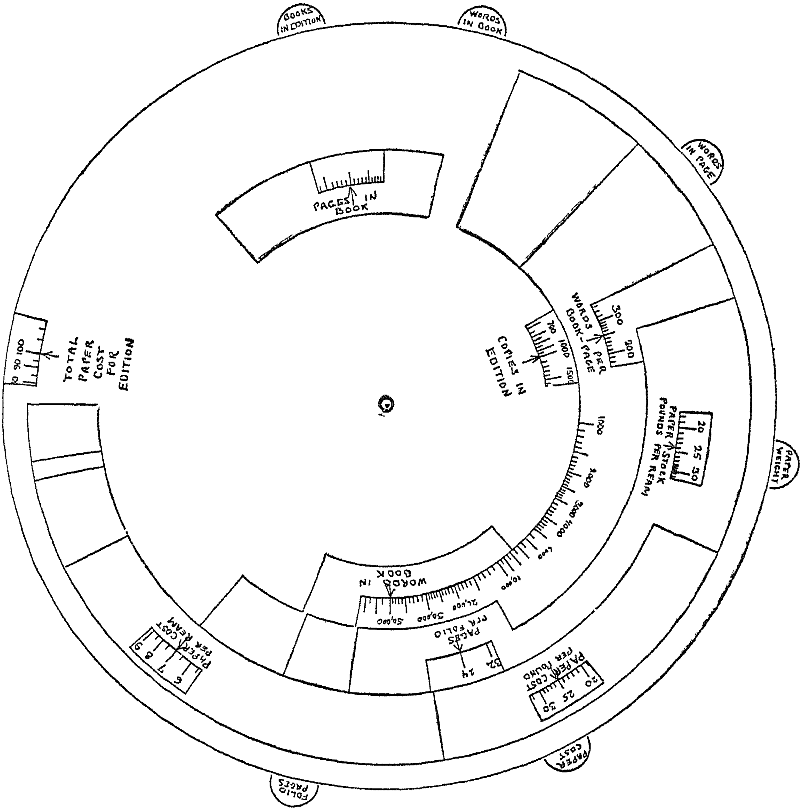
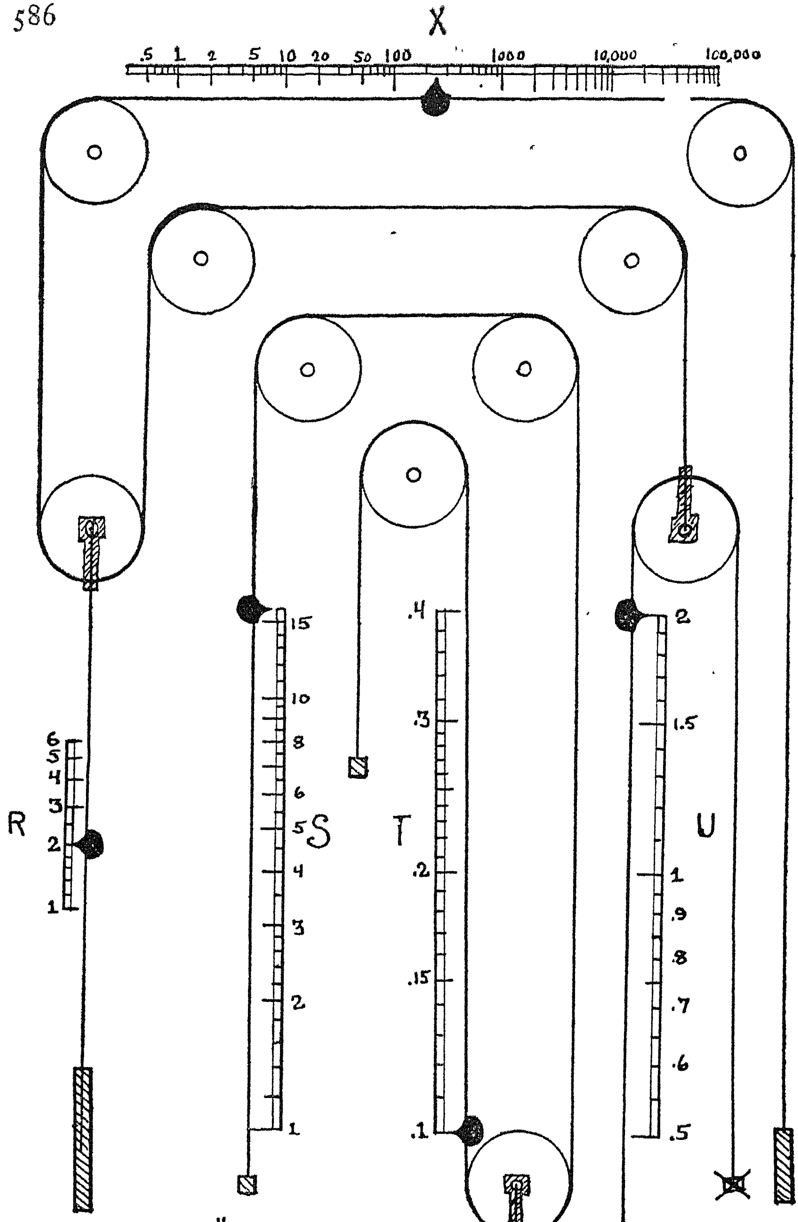


Fig. 448. The Same as the Preceding, Except that All Scales are Covered and Seen Only Through Small Open Slots or Windows.

When facilities are at hand for delicate and precise carpentry work, the straight slide-rule principle can be elaborated and developed by a series of pulley wheels with cords passing over them and connecting movable pointers along separate fixed scales. The pointers can then be adjusted for the par-

[2] The scale-moduli of slide-rules vary inversely with the coefficients (in additive rules) and exponents (in factorial, or log, rules) of the variables, and are alike as these are alike. The length of the rules or scales, therefore, varies directly with the ranges of variation of the variables (unless the moduli are unlike).

In the pulley-wheel and pointer type of slide-rule next described, the pulleys can be made with various leverages and thus modify the moduli, affording opportunity for the adaptation of the lengths of scales to any moduli, range, or length, desired.



$$X = \frac{R^4 S}{(10T)^2} = R^4 U^4$$

$$U^2 = \frac{\sqrt{S}}{10T}$$

Fig. 449. An Arrangement of Pulleys, Wheels, and Weights, by Means of which the Pointers come to Rest at the Roots of the Equation. (See note on opposite page.)

ticular readings of the variables and the final pointer will come to rest at the correct reading of the answer for the equation, the product or quotiênt of the various component factors. In such devices it is even possible to combine the different processes of addition and multiplication in one machine at the same operation, so that machines can be constructed to give the answer for any set of variables in any equation. It is hardly in the province of this book to go into the details of construction of such calculating devices. The circular slide-rules are, however, so easily constructed and sometimes such great labor-savers and time-savers that it is well to be able to make them, when occasion arises, especially adapted to your own problem.[3]

[3] It is by this time doubtless apparent that slide-rules and nomographs are clearly akin. When we have an equation with one independent variable, we have a fixed and rigid equality between it and the dependent variable, by which one is always a certain function of the other. In such cases, the chart of the equation is a chart with fixed or stationary scales. But when there are two independent variables, we can either (1) use sliding-scales, so that one of the variables can be eliminated by a proper setting, or (2) use separated scales (that is, nomographs) so that one end of the isopleth can be properly set to eliminate a variable. And over against these uses of scales, we have always a more cumbersome alternative, either for one or two independent variables, in the curve-chart.

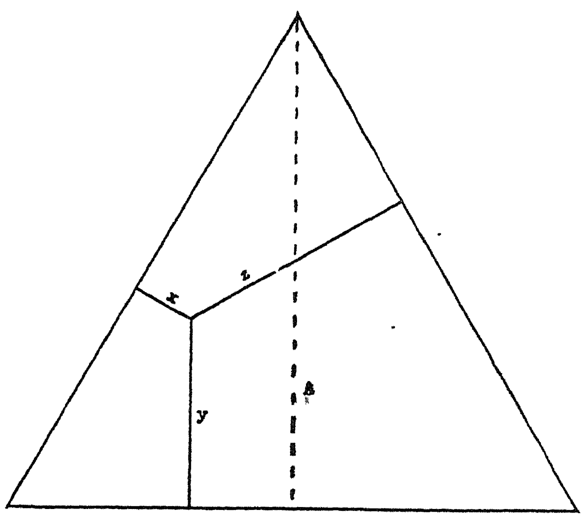
NOTE TO FIG. 449

All wheels are fixed except the three in frames, which slide up and down. The shaded bars are weights free to slide up or down, and sufficient to balance each other, so that the string remains taut and at rest wherever it is set. The small cross in the lower left-hand corner is the only fixed point to which any string is attached. Computing is done by setting any pointers at proper points and reading the answer on the remaining pointer.

Chapter XLIX

HUNDRED-PER-CENT TRIANGLES

Whenever you are dealing with problems in which three elements or parts combine to form one whole, and you are interested not in the whole but in the proportions of the three parts, the computing of the problems as well as the recording and presentation of the results can be accomplished with a chart which we may call the 100% triangle. The chart is also known as the trilinear chart and its rulings are sometimes called areal co-ordinates. It differs from the 100% bar in that it is limited to divisions of a total into three and only



$x + y + z = A$

Fig. 450.

three components, but it is similar to the 100% bar in that it does not distinguish between large and small totals, all totals

being reduced to 100% and appearing upon the chart in exactly the same size. It resembles the nomographic charts in that it is a real labor-saver in the work of computing.

Like the nomograph, the 100% triangle is based upon a trigonometric principle. The theorem in this chart is that in an equilateral triangle the sum of the three perpendiculars dropped from a point within the triangle to the sides of the triangle is constant and always equal to the altitude of the triangle. The rule is limited to equilateral triangles, and the 100% triangle is therefore always made of equilaterals.[1] The three perpendiculars bisecting each side and extending from the sides to the opposite angles of the triangle form the three "axes" of the chart. As in the calculating charts, the word "axis" is here used in the special sense of a straight line to which a scale is attached and from which ordinates are projected normally (that is, in perpendicular direction). Rulings parallel to the sides serve to project the scales across the chart in the manner of co-ordinates. A little thought will show that the 100% triangle is merely a variety of the rectilinear computing chart, in which the z-scale is calibrated so that $z=100-(x+y)$ instead of $z=x+y$, or that $100=x+y+z$. From this it follows that the equilateral triangle is not essential; any triangle can be used, but for reasons already pointed out[2] the scales are most easily projected upon the equilateral form, since then the distances along the three axes have the same real significance.

The scales used for the axes of the chart may have any range, but the customary scale is an arithmetical one, calibrated in percentages and ranging from zero to one hundred per cent. This forms the arithmetical 100% triangle. Data to be charted upon it must first be turned into percentages of the total of the three elements charted. The chart is an additive one, the three elements combining by addition to form a total. As has been said, the chart does not distinguish between the sizes of the totals, but shows them all of the same size. Nor is it possible to attempt to show relative total sizes by varying the dimensions of the equilateral triangle, as its area

[1] Obviously, any other triangular form can be used, but the intersections of the three sets of co-ordinates become less sharp and well-defined, and the scale-moduli become less easily commensurable. See Chapter XLIII.

[2] See Chapter XLIII. The 100%-triangle is most fully described in Haskell, Allan C., *How to Make and Use Graphic Charts*, Codex Book Co., New York.

increases by the square of the increase in its altitude. The general form of the equation for the chart is "100% $= x + y + z$".

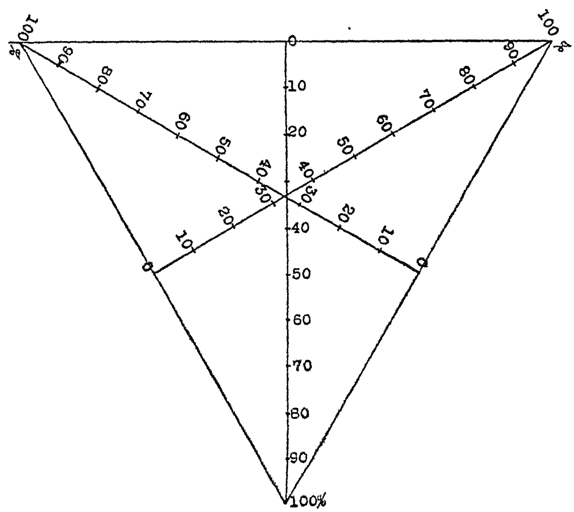


Fig. 451.

The classical example of the use of the 100% triangle is the analysis of food values in terms of calories (or heat producing units) of fats, proteins, and carbohydrates or hydrocarbons (sugars and starches). The well-balanced ration being about 20% proteids, 60% hydrocarbons and 20% fat, a point can be located on a 100% triangle at the intersection of these three ordinates and the approach of various foods and combinations of foods to this ideal easily seen. Moreover, the computing to plan a well-balanced meal can easily be done upon the chart. Thus with the two points for bread and milk plotted upon a chart, a line drawn between the two points indicates all the possible food values which can be had by mixing the two in various proportions. With equal caloric amounts of each, the point midway upon this connecting line will give the components of the combination. Foods known to lie upon the opposite side of the ideal ration-point from this mixture-point, must then be added to bring the meal nearer to the ideal, a straight connecting line again serving to show the results of combinations in all possible proportions. Such calculations can be performed upon this chart in a fraction of the time which they would require by any other method.

In business this paper can be used in innumerable cases where a total is divided into three parts. Advertizing appro-

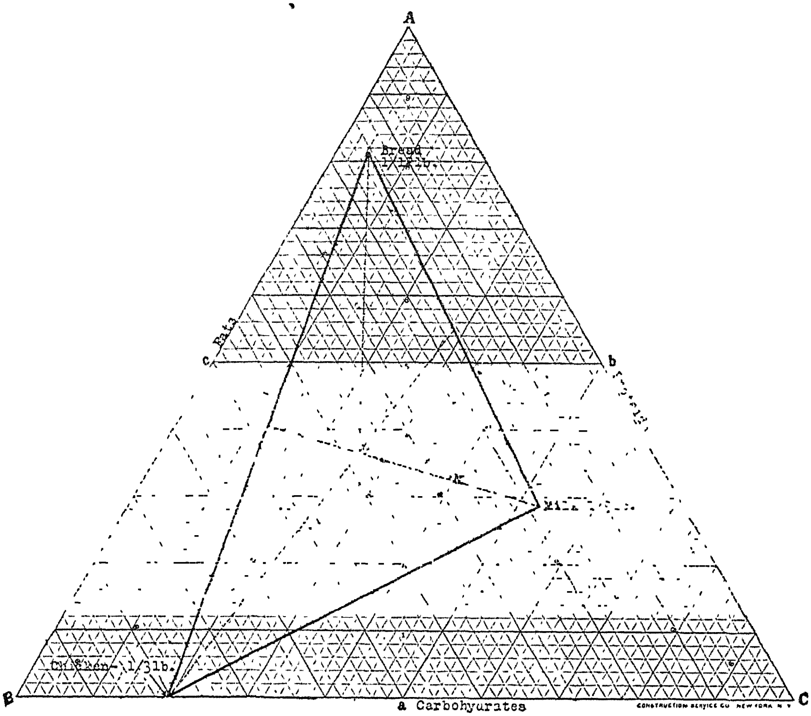
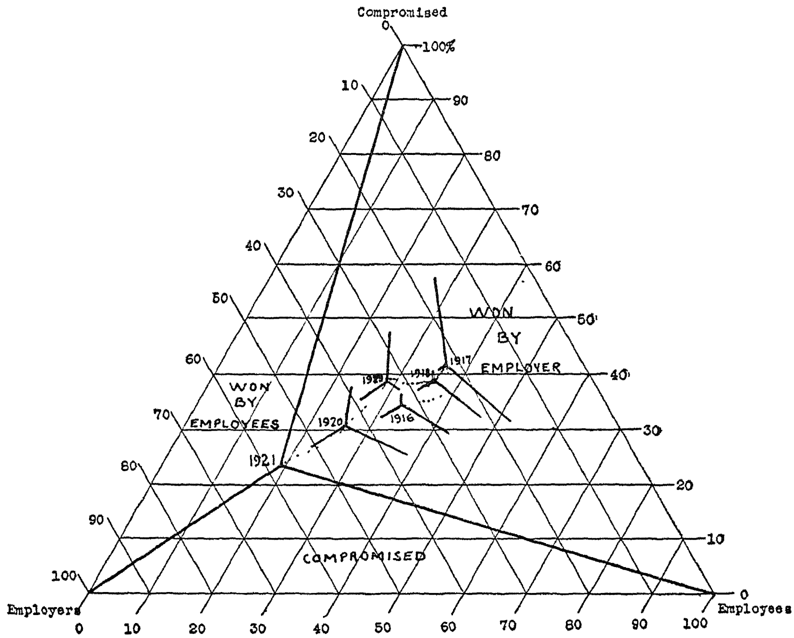


Fig. 452. The Hundred-Percent Triangle for Food Values.
The chart shows the fats, proteids, and carbohydrates (in calories) of chicken bread and milk, the full lines show the results of mixing any two of these, and the intersection of the broken lines shows the result of combinging an equal quantity of each.—*Permission of Mr. Malcolm C. Rorty.*

priations, for example, may be divided into magazine, newspaper, and outdoor advertising. Salesmen are often expected to preserve a certain proportion between their sales of large, medium, and small profit lines. Inventories may be kept in terms of raw materials, finished stock, and goods in process. Assets are often divided into current, fixed, and intangible assets; costs into payroll, materials, and overhead. In scientific work the chart has been used for the chemical analysis of mixtures of three elements. Extensive use of the chart has been made in engineering, for comparing various grades of coal as to their hydrogen, oxygen, and carbon content; and for investigating concrete mixtures and so on. In economics

the chart would seem admirably adapted to the study of projects for the joint representation of owners, workers, and



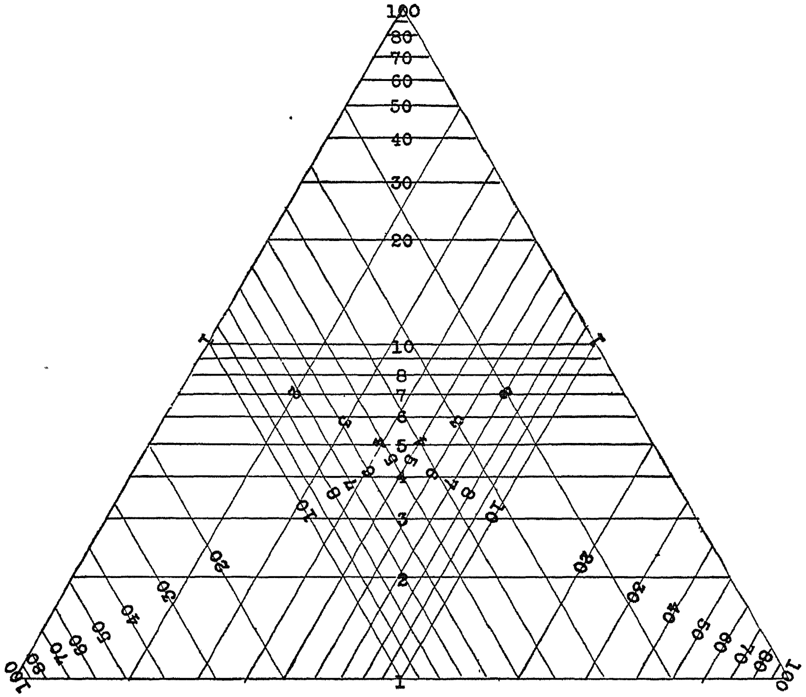
SETTLEMENT OF STRIKES
Percentages of Victories for Employers and Employees
and of Compromised Strikes
United States
1916-1921
(Arranged from Monthly Labor Review)

Fig. 453.

public in industrial disputes, or the division of earnings into wages, salaries and dividends, or the division of authority between management, workers, and stock-holders. The tripartite division is frequently called for.

By using logarithmic scales, the chart can be made factorial instead of additive and used for cases where three elements combine by multiplication to form a product. Its equation is: $\log 1 = \log x + \log y + \log z$, or $100\% = x\,y\,z$. As before, the chart will not show the size of the resulting product, it will merely show the relative proportions of the three components. The scale can be of as many logarithmic decks as desired, according to the range of variation of the components, but of course the same scale must be used for all three axes. This chart is often better when its scales are re-

calibrated to absolute quantities, the product of which is a fixed given amount. The chart can then be used for the com-

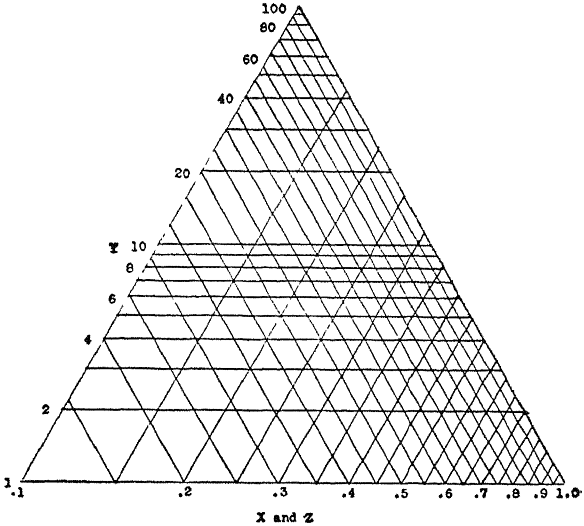


XYZ = 100

Fig. 454. The Factorial 100% Triangle.

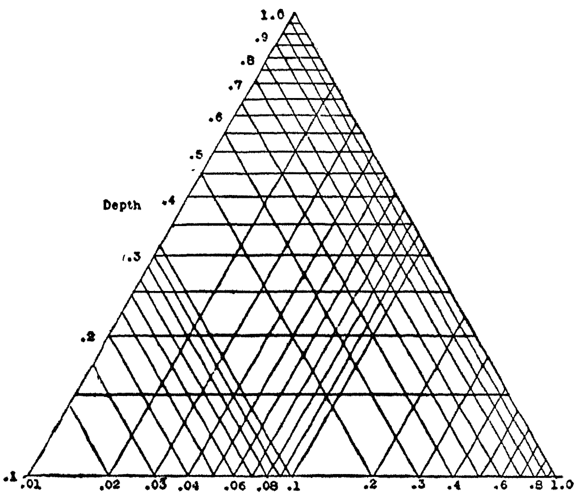
parison of the component factors in two or more equally large products. Its equation then is: log $C = \log x + \log y + \log z$, or $C = xyz$.

The logarithmic 100% triàngle has apparently never been used, but it is almost as often desirable as the arithmetic one. In engineering, for example, there are innumerable equations involving three factorial variables. A most obvious case is that of cubic measurements involving height, length, and breadth. The commercial measure of electric current is the kilowatt-hour, a product of voltage, amperage, and hours. In factory management, labor cost for a job is the resultant of the number of workmen, their average hourly wage, and their time on the job. In finance, business, and economics, similar



$$\frac{X\sqrt{Y}}{Z} = 1.0$$

Fig. 455. A Single Scale Used for Two Axes.



DIMENSIONS OF BEAMS OF EQUAL TRANSVERSE STRENGTH
(Beams of rectangular section)

$$\frac{\text{Breadth} \times (\text{Depth})^2}{\text{Length}} = 1$$

Fig. 456.

equations will be met, involving three variable factors, and susceptible to analysis by this form of chart.

With a thorough understanding of the method of the chart, it can also be used for division, by the use of reciprocals. The equation then is $\log C = \log x + \log y - \log z$, or $C = \frac{xy}{z}$. An equation of the form $C = \frac{x}{yz}$ or $C = \frac{1}{xyz}$, could equally well be shown. Without logarithms, the chart thus performs subtraction, as $100\% = X + Y - Z$ or $100\% = X - Y - Z$ or $100\% = -Y - Y - Z$. There is also another method for these cases, which does not involve recalibration of scales. It is based upon the more general geometric theorem that the algebraic sum of the perpendiculars from the sides of an equilateral triangle (or extension of the sides) to any point in the plane of

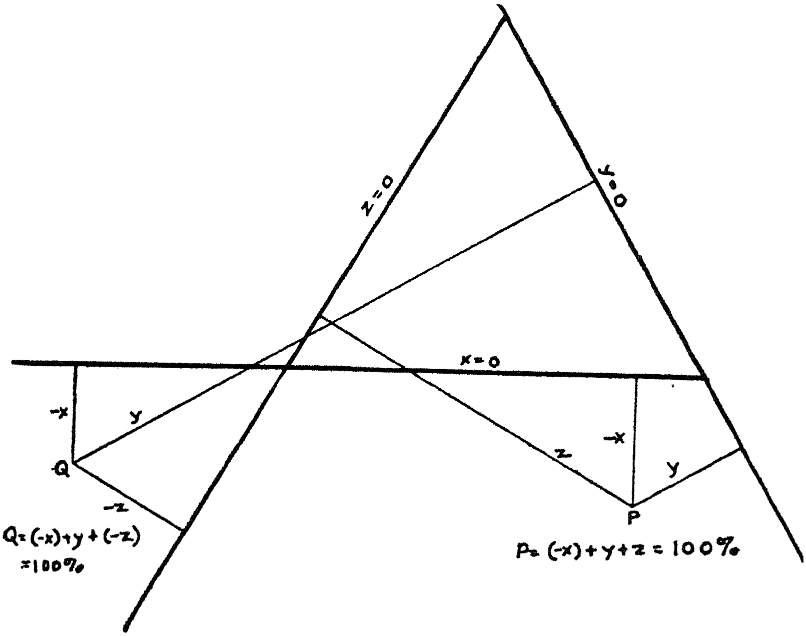
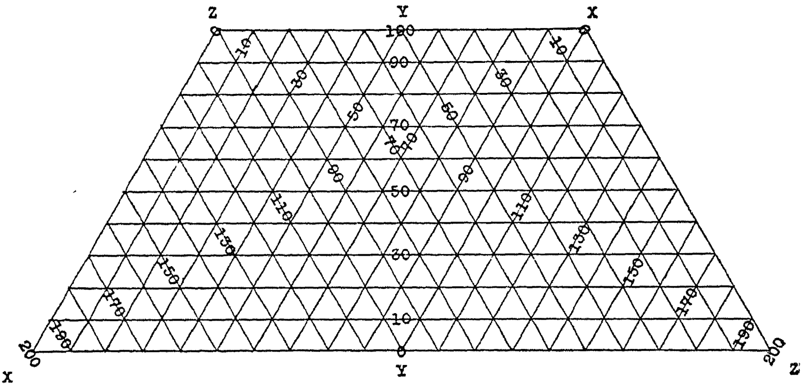


Fig. 457.

the triangle, is constant and equal to the altitude of the triangle The method involves the use of the area outside of one of the sides of the triangle, and hence. on the axis of that side, in the negative part of the scale.

In common with the 100% bar, and the 100% circle or pie-chart, the additive 100% triangle is equally significant as



X - Y + Z = 100 %

Fig. 458.

to the linear measurement and the area measurement of its parts. For the plotted point within the triangle can be connected with the three angles by three straight lines, and these

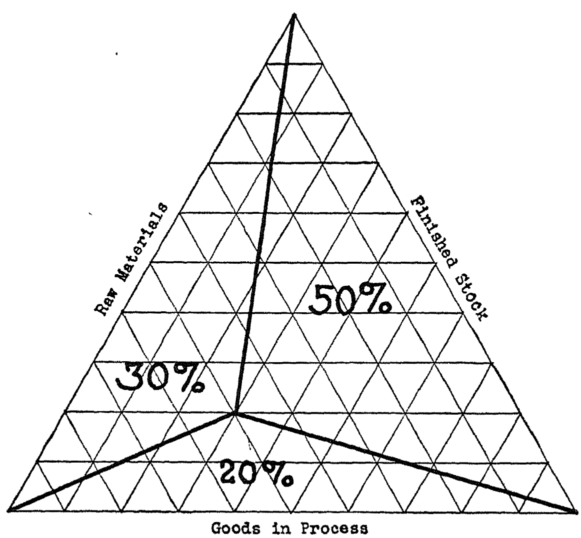


Fig. 459.

straight lines will be seen to break up the triangle into three small triangles, whose areas are in the same proportion as

their altitudes, or distances along the scales of the triangle. In this case, as in the 100% bar and pie-chart, the area of each segment is significant because only one of its dimensions varies, the other dimension being constant for the three segments or areas. Obviously, the segmentation of the 100% triangle is only useful in additive charts and only desirable for extremely popular use, when the chart is to show but one set of data, that is, one sum broken into three parts.

For both additive and factorial equations, with either absolute or relative values or units of measurement, the chart will be found strikingly illuminating and an excellent means of analysis and comparison within its somewhat narrow limitations. For it must always be remembered that the chart does not show totals, products, or other resultants. It shows only the comparative sizes of the components. If, therefore, it is the size of the resultant in which you are interested, the chart is worthless, but if it is the proportions of its three parts or factors, the chart is admirable, and is in fact, quite the clearest possible means of analysis.

PART VI. TWO- AND THREE-DIMENSION DATA

CHAPTER L

HUNDRED-PER-CENT SQUARES

We have so consistently inveighed against the use of areas to illustrate quantities that the reader will indeed be surprised at some coming retractions, however guarded and limited we may make them. But the fact is that we now propose to turn to advantage the very feature of areas which has previously been their greatest fault. Let us examine this feature closely and see how it can be done. The reader is, of course, familiar with that elementary theorem of geometry and arithmetic, which states that the number of square units of measurement in an area is a product of the linear units of measurement in its two dimensions. He therefore realizes that the variation of both linear dimensions by a given ratio results in the variation of the area itself by the square of this ratio. This has been the dangerous stumbling block in the way of using areas in charts which are intended to illustrate but a single ratio or set of ratios. To prevent the confusion and doubt which might arise in the minds of those who see our charts, it has been necessary to maintain between the areas the same ratios as exist between their linear measurements. And to maintain this identity of ratios, we have been obliged to keep one linear dimension constant wherever areas have appeared. In the 100% bar, the 100% circle, the 100% triangle, and in bar-charts, and even band charts, in short in all charts using areas up to this point, we have invariably striven to maintain one constant dimension, with this specific purpose of making variations in areas equal to variations in the one varying dimension.

We now come to data in which we wish to show simultaneously three ratios or sets of ratios, one of which is always the product of the other two. In other words, we wish to show two factors or sets of factors and their product. And to this purpose it is obvious that the area is excellently adapted,

by reason of the feature which has just been described. What was previously an obstacle to the use of areas (varying along both dimensions) now becomes an essential advantage. And as in the case of bars (that is, areas varying in one dimension only) so also in the case of areas proper (that is, areas varying along both dimensions), we find that the charts can be divided into two groups. The first group is composed of charts in which the total or whole area is a constant and is cut into segments whose sizes are of interest to us. The second group is composed of a series of separate areas, which may or may not be individually segmented, but in which the total sizes of the individual areas vary. In the section on bars, the first type was called the 100% bar; the second, a bar-chart. Similarly in this section the first type is called the 100% square or rectangle; the second, an area-chart or area-bar-chart.

Before proceeding to the separate consideration of these two types, we must call attention to a general limitation holding for all types of area-charts. These charts, as we have said, illustrate simultaneously two factors and their product, the product being shown by the area itself, the factors by linear dimensions. Now it has already been frequently pointed out that the human eye cannot so easily or precisely judge of square measures as of linear ones. Hence we must expect the illustration of the factors to be clearer and more easily evaluated by the reader than the illustration of the product. And we may say in general, therefore, that the area charts are desirable only for data in which the product itself is of less importance than one or both of the factors. Where an explicit illustration of the products is necessary, affording precise and detailed comparisons of the products, this chart does not suffice, but where the primary importance attaches to one or both of the factors, and the product is only of secondary importance, the chart will serve excellently.

Like the 100% bar, circle, or triangle, the 100% square is a device for the illustration of the parts of a total. Unlike them, however, the 100% square does not show only one classification of the parts, but shows simultaneously two independent classifications, which combine factorially to produce a great many small parts. The data for the 100% square, therefore, consists of two interlocking or mutually crossing and subdividing classifications of the parts of a whole. Whatever the absolute value of this whole may be, its relative value is

100%, and the chart is ordinarily calibrated in percentages. If the data is not already in percentages, it can easily be turned into percentages to facilitate charting. The important

	Total	Proprietors Managers Officials	Clerks and kindred Workers	Skilled Workers	Semi-skilled Workers	Laborers and Servants
Total	41,609,192	11,165,536	5,638,144	4,914,651	6,384,567	13,506,294
Agriculture, forestry and animal husbandry	10,951,074	5,501,742	---	---	---	5,449,332
Extraction of minerals	1,090,854	28,610	---	---	509,041	553,203
Manufacturing and mechanical industries	12,812,701	660,622	---	4,689,126	4,247,232	3,215,721
Transportation	3,066,305	206,352	455,305	225,525	648,419	1,530,704
Trade	4,244,354	1,615,823	2,062,884	---	355,205	210,442
Public service (not elsewhere classified)	771,120	658,351	---	---	---	112,769
Professional service	2,152,464	2,128,787	---	---	23,677	---
Domestic and personal service	3,400,365	365,249	---	---	600,993	2,434,123
Clerical occupations	3,119,955	---	3,119,955	---	---	---

OCCUPATIONS OF THE GAINFULLY EMPLOYED POPULATION
(10 years of age and over)
The United States
1920
(Source:- U. S. Bureau of Labor Statistics)

Fig. 460. The Original Data for a 100% Square.

thing about the data is that it should be clearly arranged in a tabulation or table, with the items of one classification listed down an edge of the table as the stubs of the table, and the

	Total	Proprietors Managers Officials	Clerks and kindred Workers	Skilled Workers	Semi-skilled Workers	Laborers and Servants
Total	100.00	26.84	13.54	11.83	15.33	32.46
Agriculture, forestry and animal husbandry	26.31	13.22	---	---	---	13.09
Extraction of minerals	2.62	0.07	---	---	1.22	1.33
Manufacturing and mechanical industries	30.82	1.59	---	11.29	10.20	7.74
Transportation	7.37	.50	1.09	.54	1.56	3.68
Trade	10.20	3.88	4.96	---	.85	.51
Public service (not elsewhere classified)	1.85	1.58	---	---	---	.27
Professional service	5.18	5.12	---	---	.06	---
Domestic and personal service	8.16	.88	---	---	1.44	5.84
Clerical occupations	7.49	---	7.49	---	---	---

OCCUPATIONS OF THE GAINFULLY EMPLOYED POPULATION
(10 years of age and over)
The United States
1920
(Source:- U. S. Bureau of Labor Statistics)
(percentages)

Fig. 461. In this Form the Data is not Chartable.

items of the other classification listed across the top of the table as its column headings. In the body of the table, at the

intersections of columns and rows, are placed the detailed figures which correspond to both classifications.

In turning the absolute detail figures into percentages of the total, we have, of course, no trouble, merely dividing each figure by the figure for the total. But in this form, the data is no longer factorial, the detailed figures not being products of factors which are known, and therefore not being amenable to charting by areas. To draw areas we must know the factors which will be plotted as the linear measures along the two dimensions of the area. It is therefore of no use to us to turn the absolute values into direct percentages of the grand total. Instead, we turn the sub-totals for each row or column into percentages of the grand total, and then turn the detail

	Total		Proprietors Managers Officials	Clerks and kindred Workers	Skilled Workers	Semi-skilled Workers	Laborers and Servants
	Vertically	Across					
Total	100.0	100.0	26.84	13.54	11.83	15.33	32.46
Agriculture, forestry and animal husbandry	26.31	100.0	50.3	---	---	---	49.7
Extraction of minerals	2.62	100.0	2.6	---	---	46.7	50.7
Manufacturing and mechanical industries	30.82	100.0	5.2	---	36.6	33.1	25.1
Transportation	7.37	100.0	6.7	14.8	7.4	21.1	50.0
Trade	10.20	100.0	38.1	48.6	---	8.4	4.9
Public service (not elsewhere classified)	1.85	100.0	85.4	---	---	---	14.6
Professional service	5.18	100.0	98.9	---	---	1.1	---
Domestic and personal service	8.16	100.0	10.7	---	---	17.7	71.6
Clerical occupations	7.49	100.0	---	100.0	---	---	---

OCCUPATIONS OF THE GAINFULLY EMPLOYED POPULATION
(10 years of age and over)
The United States
1920
(Source:- U. S. Bureau of Labor Statistics)
(percentages)

Fig. 462. Here Each Row Totals 100%.

figures in the body of the table into percentages of the sub-totals for the rows or columns in which the detail figures occur. Thus we make the detail figures in themselves products, that is, percentages of percentages.

In this step we come to a choice between turning the detailed figures into percentages of the sub-totals for the columns or into percentages of the sub-totals for the rows, in which they occur. One or the other must be used, that is, either the sum of the detail percentages in each column must add up to 100%, or the sum of the detail percentages in each row must add up to 100%. We cannot expect that addition up and down by columns, and addition across by rows, will both

give 100% in the same table. We must, therefore, make a distinction between the primary classification, in which the sub-totals are percentages of the grand total, and the sec-

	Total	Proprietors Managers Officials	Clerks and kindred Workers	Skilled Workers	Semi-skilled Workers	Laborers and Servants
Total (across	100.00	26.84	13.54	11.83	15.33	32.46
Total (vertically	100.00	100.00	100.00	100.00	100.00	100.00
Agriculture, forestry and animal husbandry	26.31	49.2	---	---	---	40.3
Extraction of minerals	2.62	0.3	---	---	8.0	4.1
Manufacturing and mechanical industries	30.82	5.9	---	95.4	66.5	23.8
Transportation	7.37	1.9	8.1	4.6	10.1	11.3
Trade	10.20	14.4	36.6	---	5.6	1.6
Public service (not elsewhere classified)	1.85	5.9	---	---	---	0.9
Professional service	5.18	19.1	---	---	0.4	---
Domestic and personal service	8.16	3.3	---	---	9.4	18.0
Clerical occupations	7.49	---	55.3	---	---	---

OCCUPATIONS OF THE GAINFULLY EMPLOYED POPULATION
(10 years of age and over)
The United States
1920
(Source:- U. S. Bureau of Labor Statistics)
(percentages)

Fig. 463. Here Each Column Totals 100%.

ondary classification, in which the detail figures are percentages of the sub-totals. It is not a matter of importance how we place these classifications, but as a general rule in tables, the primary classification should be listed in the column-headings and the secondary classification in the stubs, to facilitate checking up on the computing. Where one classification is much more lengthy than the other, it is of course generally more convenient to arrange the longer classification in the stubs and the shorter one in the column-headings. In charting, the rule is generally reversed, the primary classification being shown along the vertical axis of the chart and the secondary one being shown horizontally. The chart itself always shows very clearly which classification has been made of primary importance and which of secondary importance. It often happens that both classifications appear on their merits to be equally important, but it is nevertheless necessary that the distinction be made and the data must be prepared in the form described before the chart can be made.

The chart is made by laying out a square with co-ordinate rulings. Along both axes of the square, that is, along its vertical

and horizontal edges, a scale is marked off in percentages from 0% to 100%. Arithmetic projection of the scales is used only in area charts; and in the usual form, that is, in the truly square-shaped chart, both scales are identical. The primary classification of the grand total is laid off upon the vertical scale by means of horizontal lines extended across the chart to form layers of a uniform length but of varying widths or depths.

At this stage, the chart reminds us of a 100% bar turned on end and made very short and thick, for the chart bears as yet

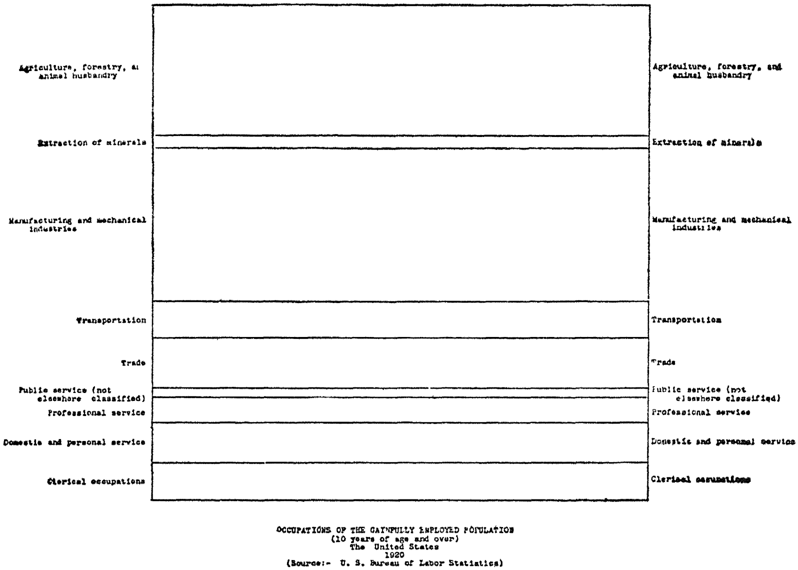


Fig. 464. The Primary Division Alone Plotted from Fig. 462.

but one classification and its segments or layers show, both by their depth and their areas, the figures for this one classification of the parts of the grand total. The layers, however, need not be distinguished by colors or shadings, for as will shortly be seen, they will be sufficiently distinguished by the markings of the secondary classification. The next step is to enter these secondary classifications. Each layer is now treated as a separate 100% bar and divided up as indicated by the detail figures in the body of the table of data. Notice that each layer is separately segmented, by vertical dividing lines which may or may not vary in their positions from layer

to layer. These segments of the layers are now colored or shaded to distinguish then, and the chart is complete. A key

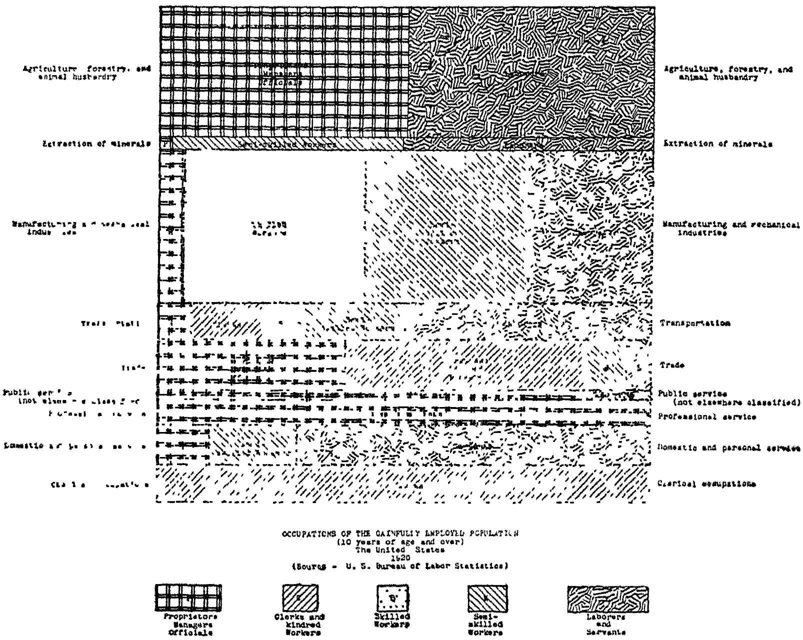


Fig. 465. The Completed Square.

to the shadings should be added, to guide the reader of the chart. And it will be seen that the area of each shaded segment of a layer is to the area of the entire square, as the absolute value of each detail figure in the body of the table of original data is to the absolute value of the grand total.

Various modifications of the 100% square are sometimes useful. The scale of the primary classification may be calibrated in absolute values instead of percentages. In this case, the square-shaped outline of the entire chart is often discarded, and the chart made rectangular, thus becoming a 100% rectangle. This is often done where the detail of secondary sub-divisions is great and the areas of segments would be so small as to be ill shown except by enlarging one of the scales. Either scale may be enlarged in this way, according to the nature of the data. It is also possible to project the primary classification upon the horizontal axis and the secondary one vertically. In this form, the chart strongly resembles the

staircase relative band curve chart with ordinates at irregular intervals. The rectilinear lines used to segment the layers

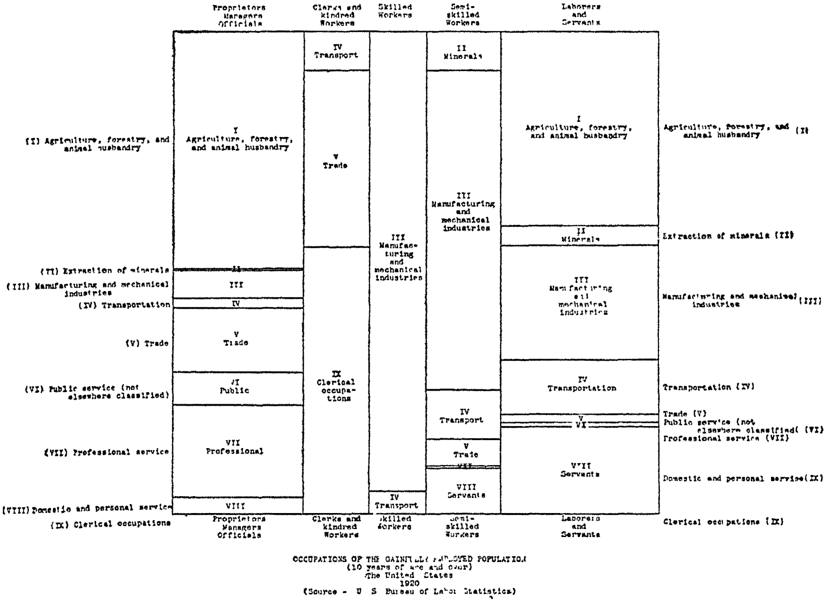


Fig. 466. Here the Primary Division is the Horizontal One, Plotted from Fig. 463.

form staircase curves separating the colored or shaded bands which form the sub-totals according to the secondary classification.

The 100% square or rectangle becomes identical with the relative band-chart (with staircased curved) when the primary classification has a numerical basis, forms an ordered mathematical series; and can be called a variable. So that the whole subject of relative band curves charts may indeed be considered a detail of the 100% area. Frequency series fall particularly well under either head. And because the primary classification can be shown upon either axis of the 100% area, it often happens that what is really only a relative band frequency curve chart turned upon edge, seems at first to be a new kind of chart. When smoothed curves are substituted for the staircase curves, by connecting the mid-points of the secondary segmenting lines, of the 100% area, the highly descriptive name of "marble-cake chart" has sometimes been used for the resulting picture. This is a very interesting form

of the 100% area. Obviously slight errors creep into the area-representation in this form, but when the change in the sec-

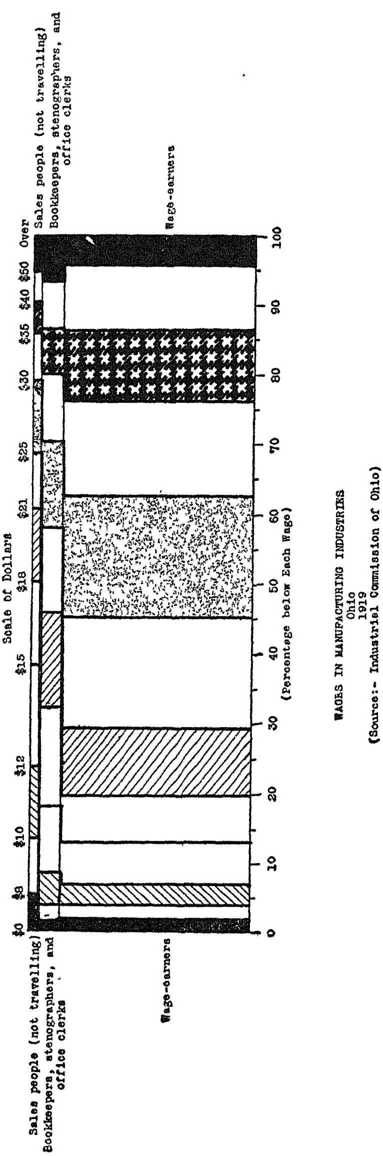


Fig. 467. A 100% Rectangle.

ondary classification is really gradual and not abrupt, the chart has gained in interpretative powers. It is commonly

useful in the analysis of the component parts of a frequency series, the plotting of the independent variable (or primary classification) along the y-axis affording greater popular appeal through the coincidence of increasing numerical values and rising position upon the chart.

The 100% square can be used for data which is not classified by, or dependent upon an ordered numerical series and in which there really are no two interlocking schemes of classification, but merely one independent variable. In such cases scales are useless on the chart and the chart itself is wholly pictorial.

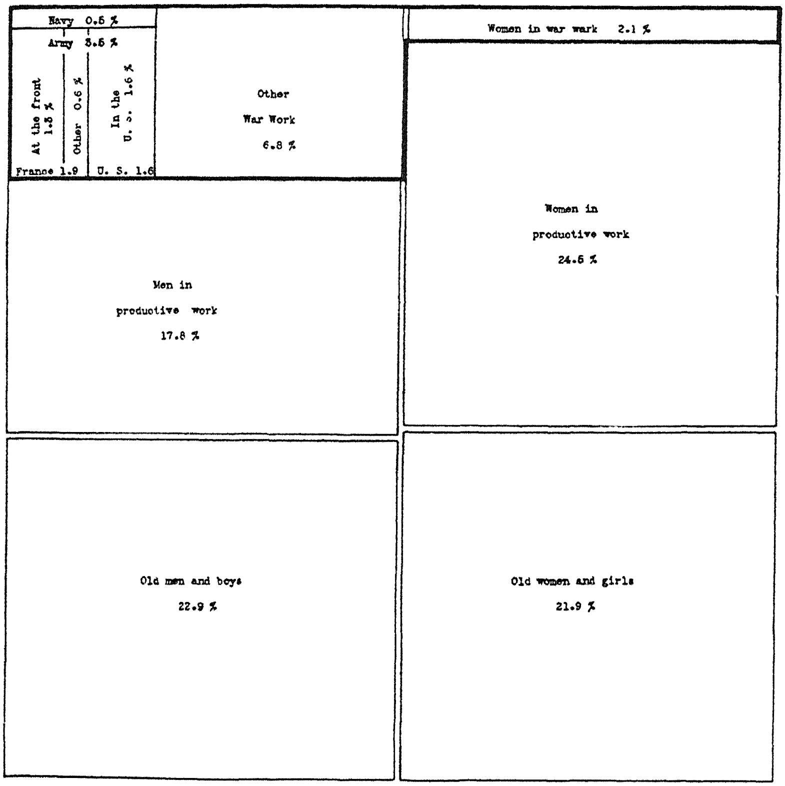


Fig. 468. A 100% Square.
Showing Wartime occupations of the population, U. S., 1918, according to official estimates; taken from the Annual Report of the Secretary of War for 1919. Total population, 105,000,000.

It has already been said that area charts must be projected arithmetically upon both axes, for the reason that only upon

this projection does the area itself illustrate the product of the linear dimensions. When a "marble-cake chart," for example, is drawn with its primary classification upon a logarithmically projected scale, the areas upon the chart lose their significance, and the chart itself really becomes merely a chart of frequency curves. The logarithmic projection may be necessary to straighten the curves or to show parts of the data in sufficient detail, but great care must be exercised that the reader of the chart should not, under these circumstances, attach the slightest importance or significance to areas.

The student who has noted how the 100% bar is particularly adapted to showing the division of a whole into two

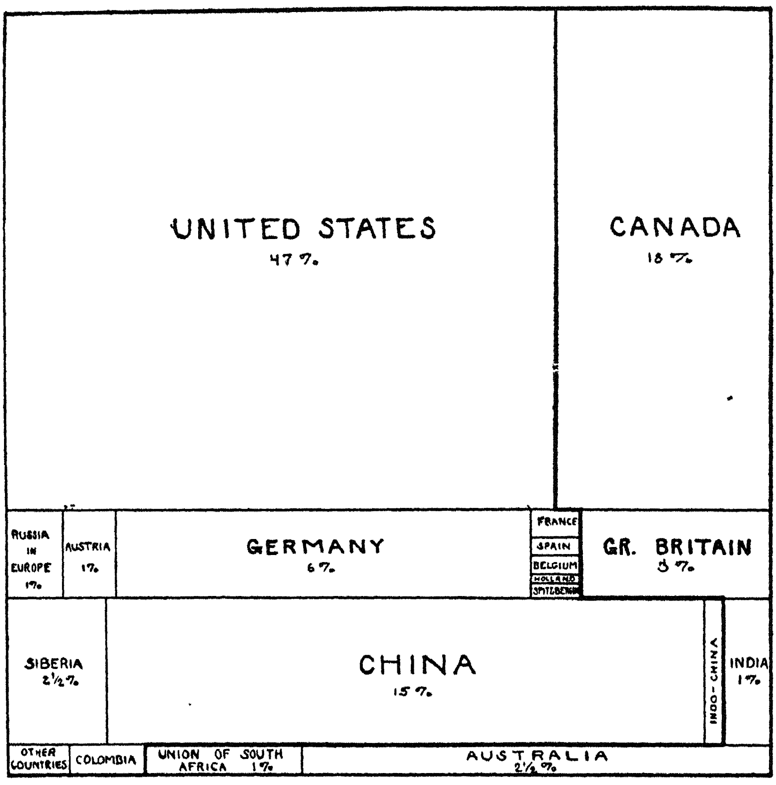


Fig. 460.
Showing the estimated unmined coal supplies in 1920.

parts (though it can show any number of parts) and how the 100% triangle is particularly adapted to showing the division

of a whole into three parts, may now be asking himself for a chart form which will show conveniently the division of a whole into four or five parts. In this case he will possibly have use for a hundred per cent square in which the four or five segments are indicated only by points and arrows or short lines. Thus if we are dealing with the sales of the four lines of a company in many different sales districts, we can combine them into two groups of two each and plotting each group as layers, we can indicate the division lines in the layers by short lines from the layer-division line only for its distance between

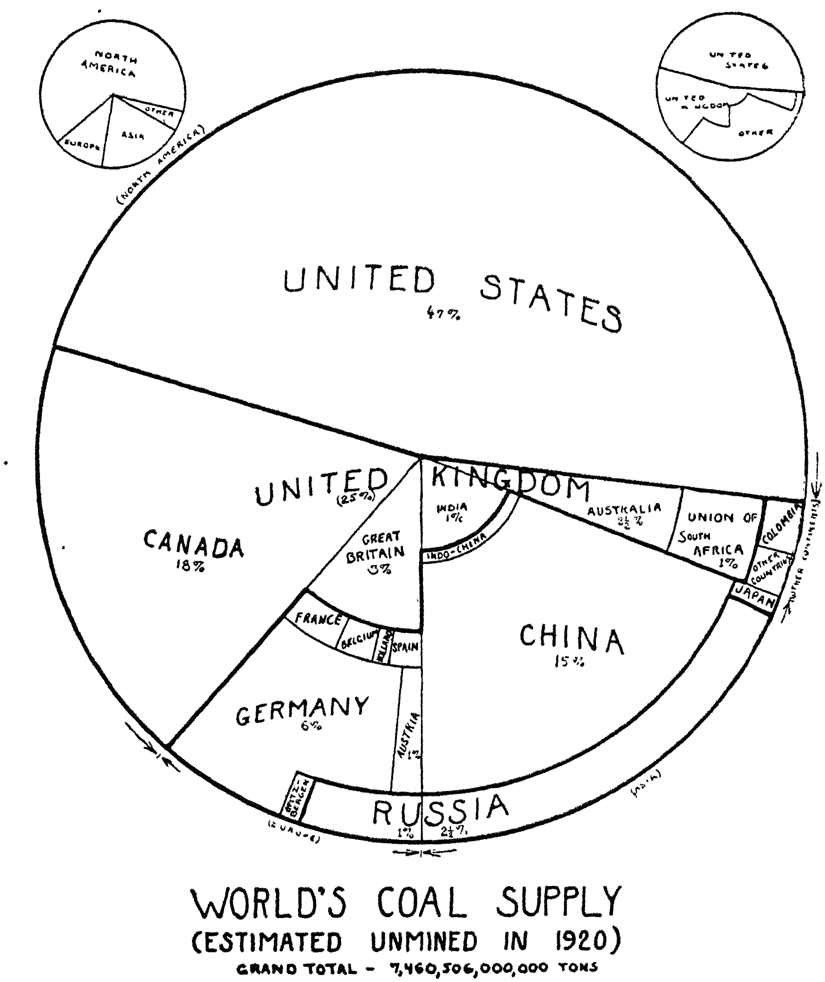
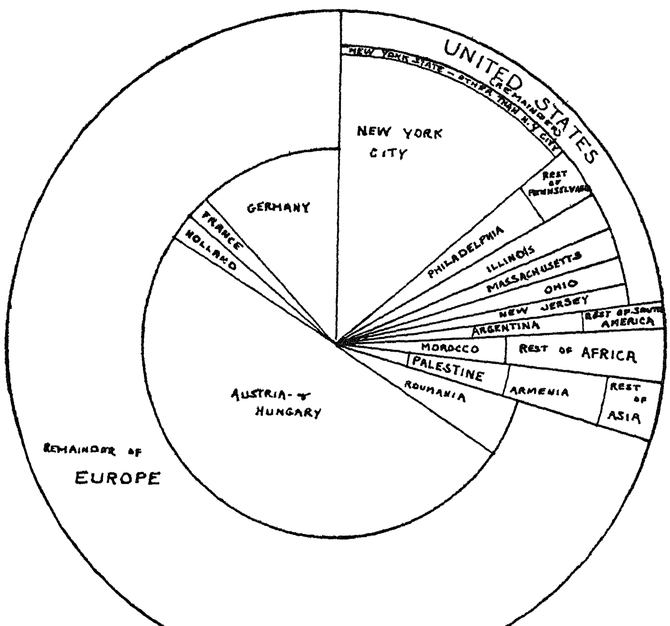


Fig. 470. Same as the Last in Circular Form.

these points. Such a chart reminds us of the famous "swastika" pattern. The chart is not of much general value, but would so abbreviate the rulings of the 100% square that many 100% squares could be superimposed or combined upon one chart with a visible record for all. Of course, the sizes of the areas would only have relative significance here, as the total areas for all grand totals would be the same regardless of their absolute values.

Closely related to the 100% square is a special type of 100% circle or pie-chart which has recently come into vogue. By means of an elaborate method of segmentation, small percentages and complicated groupings of parts can be shown without difficulty. For it is obvious that by the simple method



JEWISH POPULATION OF WORLD
1920 ESTIMATES
GRAND TOTAL = 15,000,000

Fig. 471.

of segmentation of the pie-chart, in which each segment or part of the circular area extends from center to circumference, the small segments or parts become long thin attentuated

areas, which are not easily labelled. The more elaborate method breaks the circular area into concentric rings, the width of each ring being particularly calculated to fit a particular segment in the circle, and to result in significant areas within the segment inside and outside of the ring. The making of a chart of this kind is not as easy a matter as with the simpler method, when all segments extend from center to circumference. For virtually each angular segment, or slice of the pie-chart is subjected to crosswise segmentation, and the ring-like division lines, or arcs, must be placed at distances from the center which correspond, not to the ratio of the parts to the whole of the segment, but to the square root of that ratio. The calculating is not easy. But the chart has very definite advantages for detailed and minute data when it is desirable to show several groupings simultaneously. There is little to recommend it for purposes of precise and comparative study, but for popular and unscientific purposes it is much favored.

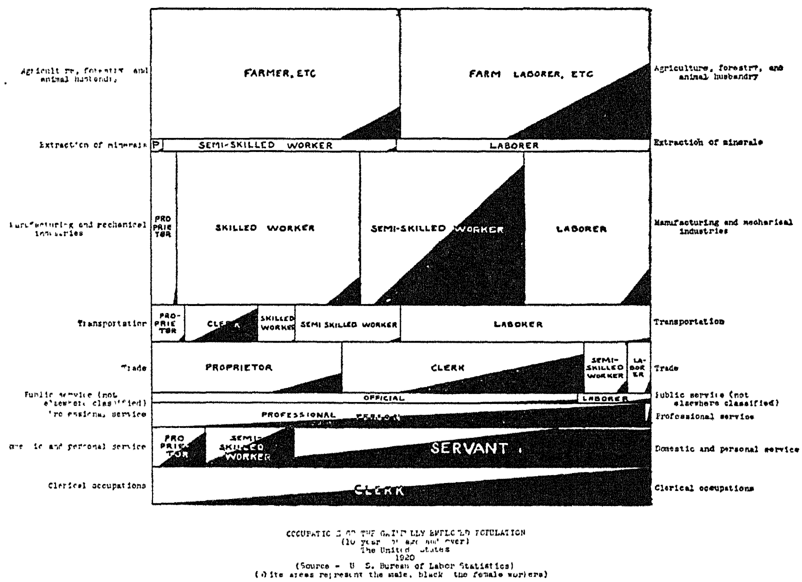


Fig. 472. A Third Classification has been Added Here by Diagonal Divisions and Shadings, Showing Sex.

The 100% square or rectangle, and its many variations, are adaptable to a wide variety of uses. No set rules can be laid down to limit the various ways in which it may be applied.

But a very careful study of the results should always be made, to ascertain that one of the simpler methods in which areas have no especial significance would not, after all, have produced more simple and forceful results. The danger is not that the ingenious chart-maker will fail to utilize all the possible significant features of the compound area chart, but that he will utilize too many of them, overcrowding his chart with complex details. The simpler, the better, both for research and publicity. And as area charts are more generally popular in their appeal, simplicity is a cardinal virtue.

Chapter LI

AREA-BAR-CHARTS

It has been already laid down as a general rule that area charts (that is charts in which both dimensions of a charted area vary) are useful only when the data represented by the area is of less importance than the data of its factors. For the area chart is based upon the geometrical theorem that the number of units of measurement in a rectangular area is equal to the product of the linear units of measurement along its two sides or dimensions. From this it follows that we can always show a numerical value by an area whenever we can break that numerical value into two factors and can plot these two factors as the two dimensions of the area.

In some cases the presence of two factors in the data or the fact that the data is the product of two factors, is so obvious as to be self-apparent. Thus the floor space of a room is obviously the product of its length and its breadth, and a chart of the room showing its dimensions and resulting area, could be constructed by the veriest novice. But should we come to compare a number of such rooms, it would be a real question whether to show the dimensions of the rooms or to show only their total areas, that is, whether to use an area-bar chart or an ordinary bar-chart. If the figure for total areas (or square feet) is more important, we must drop these variable-area charts and present the data of square feet along one dimension only by a bar-chart. If, on the other hand, it is the shape or dimensions of the rooms in which we are more interested, then of course we should adhere to the area diagram and let the reader rely upon guesswork or upon appended data for the total area. When both aspects of the data are important, it would indeed be best of all to use both methods, outlining the shape of the room by small area diagrams and showing their comparative sizes by a bar-chart. This example of data of square measurements of a physical area excellently

illustrates the fact that even data of the most obviously two-dimensional nature is, so far as the product or resultant is concerned, best shown by a one dimension chart.

On the other hand, data which seems most clearly to be one-dimensional in its nature, can always, if you desire, be broken up into two factorial parts. When this is done and you regard the factorial parts or factors of the data as more important than the data itself, you can then proceed to show these factors with their products by an area-bar chart. This breaking up into factors, it may be remarked, can always be obtained by a process of division. Thus the sales of our company in various States may be divided by population of these States, and so the per capita sales will be obtained. The same total sales in each State might also be divided by the number of dealers in each State and so the sales per dealer be obtained. Or these total sales might be divided by the similar total sales of the previous year, and so the percentage of increase be obtained. In short, the most palpably one-dimensional data may, by the process of division, be turned into factorial two-dimensional data and shown by the area chart.

Making the chart, it is neither desirable nor commonly feasible to place the zero line of both axes of all the areas together, for this would require that they be superimposed upon each other. The result would be the same as if, in making the multiple bar-chart, we had superimposed the corresponding bars, for each lower bar would be at least partially hidden by the upper one and if the upper one were at any time longer or larger than the lower one, the lower bar would of course be entirely hidden. This method of superimposition is occasionally used in area charts when the difference between the compared areas along both dimensions is very great. We then have the effect of squares or rectangles within squares or rectangles, the inner one being placed at one corner of the outer one. The reader must then be carefully warned that the larger area includes the smaller one and is not alone composed of its visible portions outside of the smaller one. In general, the method is unsatisfactory and to be avoided.

The proper method of showing areas to be compared is to arrange them side by side, so that along one of the dimensions only, the area will have a common zero line or base-line. The result then closely resembles a bar-chart, its only distinction being that the bars which form the areas are not of a

constant width, as in the bar-chart, but are of varying widths, the variations in width showing the second factor in the data. And it will be seen that both the varying widths and the resulting areas are of secondary importance, serving to give the reader of the chart a general impression of the relative importance of the items which are described by the various lengths of the area-bars. For as has been repeatedly pointed out, the reader will have difficulty in precisely comparing areas of different sizes, and it is obvious that he will also be unable to gain exact impressions of the various widths. The most that can be said for the chart is that it gives him a precise knowledge of statistics of one factor in the data (as shown by the lengths of the bar areas)—in this the chart has all the virtues of a bar-chart—and that it also gives

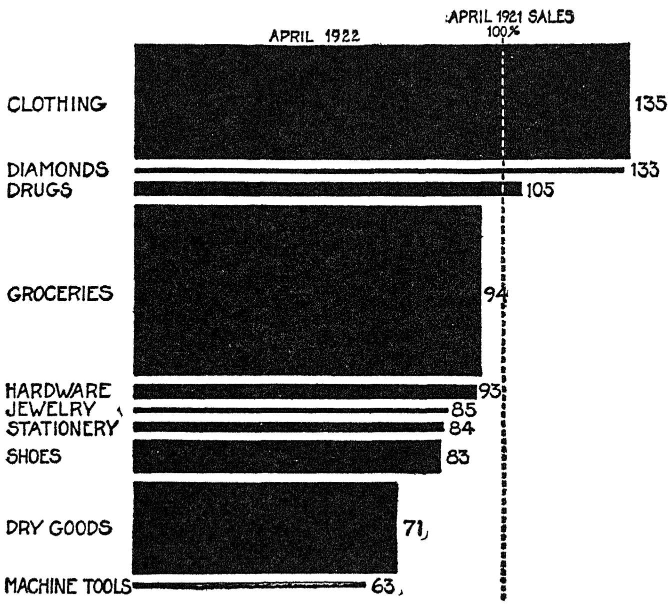
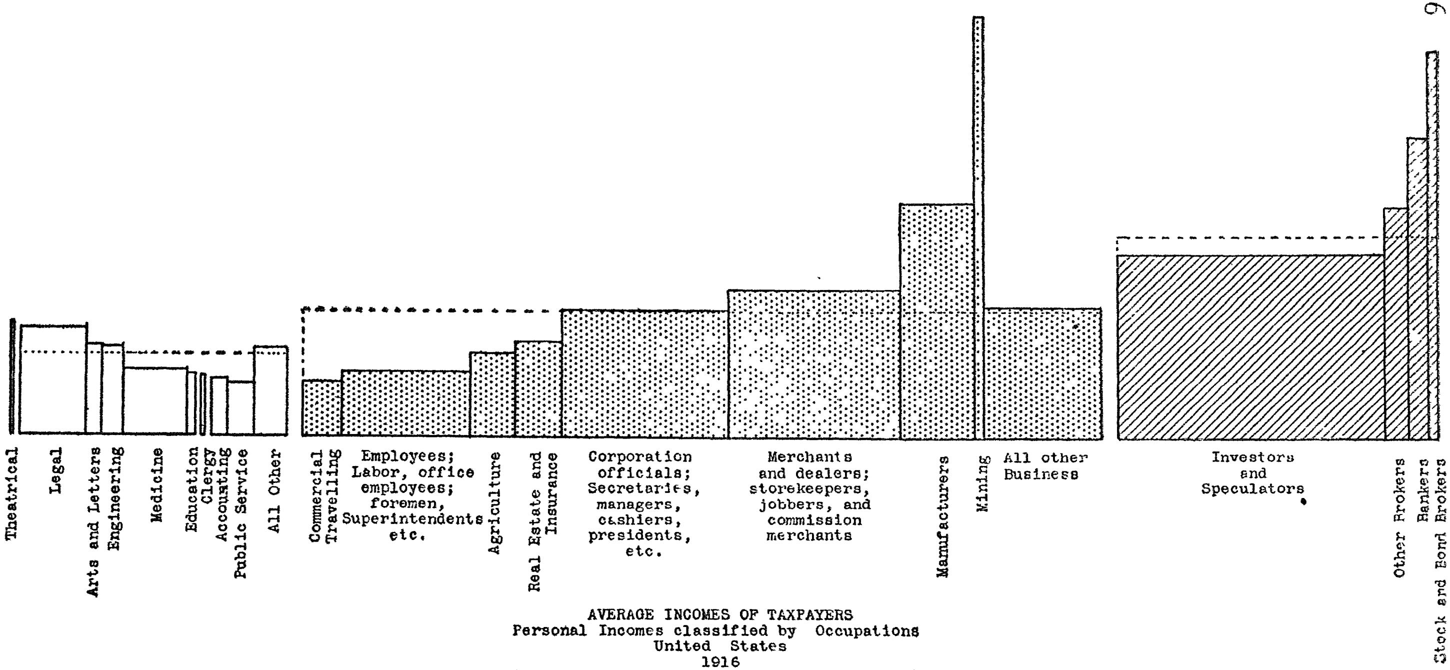


Fig. 473. A Simple and Excellent Area Bar-chart.

Sales of wholesale concerns in the Second Federal Reserve District in April, 1922, compared with their sales in April, 1921. Width of bars indicates relative amount of goods sold.—*Permission of Mr. Carl Snyder.*

him a general impression of the other factor and of the resulting product of the two factors—in this the chart is an improvement upon the bar-chart.



AVERAGE INCOMES OF TAXPAYERS
Personal Incomes classified by Occupations
United States
1916
(Source:- Commissioner of Internal Revenue)

Fig. 474. Vertical Area-Bars.

The natural arrangement of the area bars would appear to be in a column down the page. The statistics and labels or items would then also be placed in columns to the left of the bar area, the whole chart closely resembling the bar-chart. To this chart we may give the name of "area-bar-chart." The chart is dignified, sound, and extremely illuminating. It requires but little more labor than the ordinary bar-chart, for the minor importance of the second factor and the resulting product, shown by widths and areas of the bars, makes it unnecessary for these to be plotted with extreme accuracy. If a scale for the widths be chosen small enough, the data and labels or items appended to the chart can be entered at fairly uniform distances down the chart, making for a very presentable appearance. The small scale upon which the widths of the bars are plotted deters the reader from attempting precisely to estimate the secondary or less important factor and the area, while it nevertheless gives him an excellent idea of the relative importance of the primary data presented in bar-chart form. The chart is a direct outgrowth of the simple bar-chart, and in its proper place a decided improvement.

By far the more popular form of area-bar-charts however, is the modification of the pipe-organ or vertical bar-chart. To this form, the name of "sky-line chart" is sometimes given. In the sky-line chart, however, it is customary to give to the widths of the bars, a somewhat larger scale, showing their variations with more precision and emphasizing differences in areas. Partly on account of the greater widths and partly for the increased spectacular effect (which is always desired in popular charts) the areas are placed in direct contact with each other, being strung out across the page in vivid resemblance, let us say, to such a silhouette as the sky-line of lower New York City as it is first seen by a visitor from abroad. In this chart distinctive shadings or color tints are often desirable to distinguish the areas because of their close contact with each other. For the same reason very narrow bars which could only be shown by thin vertical lines are better shown with narrow separating margins between the lines or in the form of the previously described area-bar-charts.

In the choice of shadings in these charts, as elsewhere, where shadings are used, care must be taken to avoid optical illusions, produced by bringing together shadings of different color density, for it will be found that two equal areas will

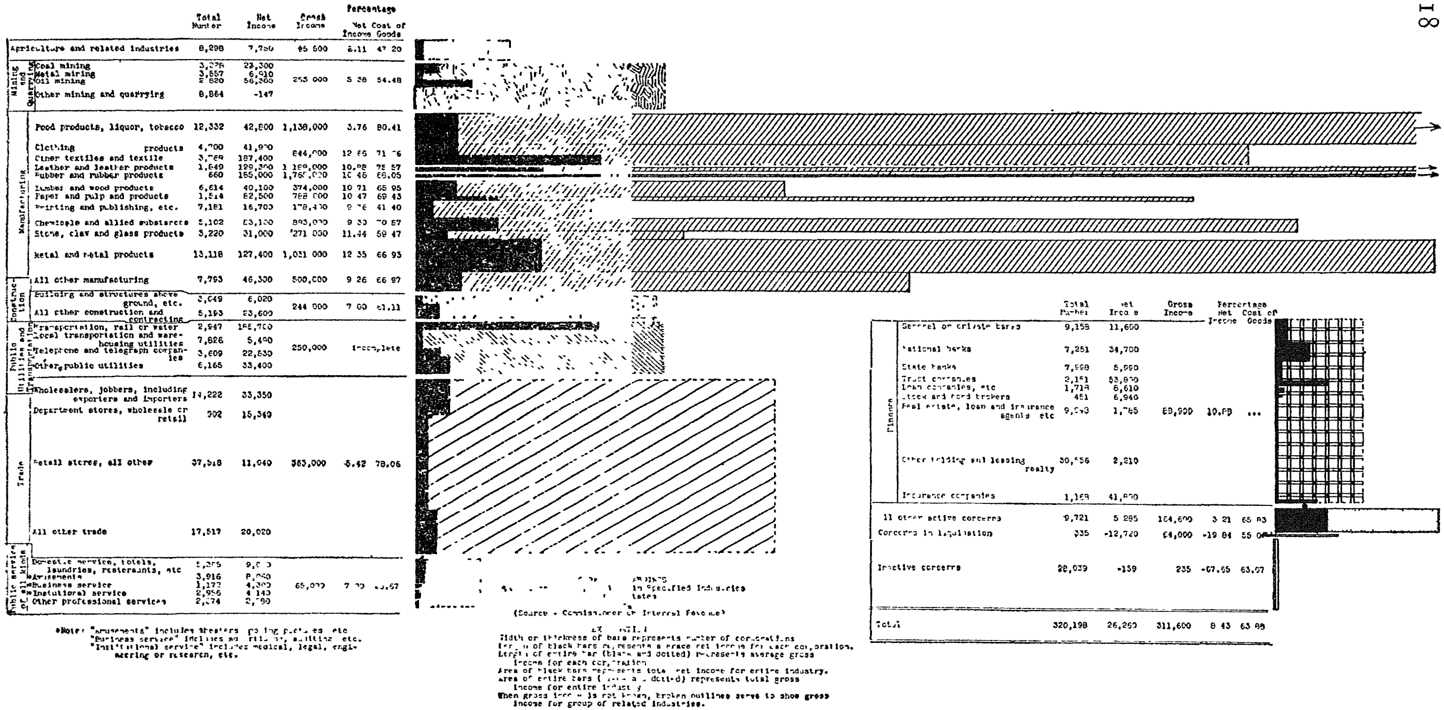


Fig. 475. Compound Area-Bars.

The shaded bars show gross earnings and the black bars net earnings of corporations.

appear unequal if one is much more densely shaded than the other. In the sky-line chart care must also be used in the entering of data and item labels, for the same difficulties of typography which were met with in the pipe-organ chart may be encountered here and the same considerations (discussed in the chapter on pipe-organ charts) in general, apply.

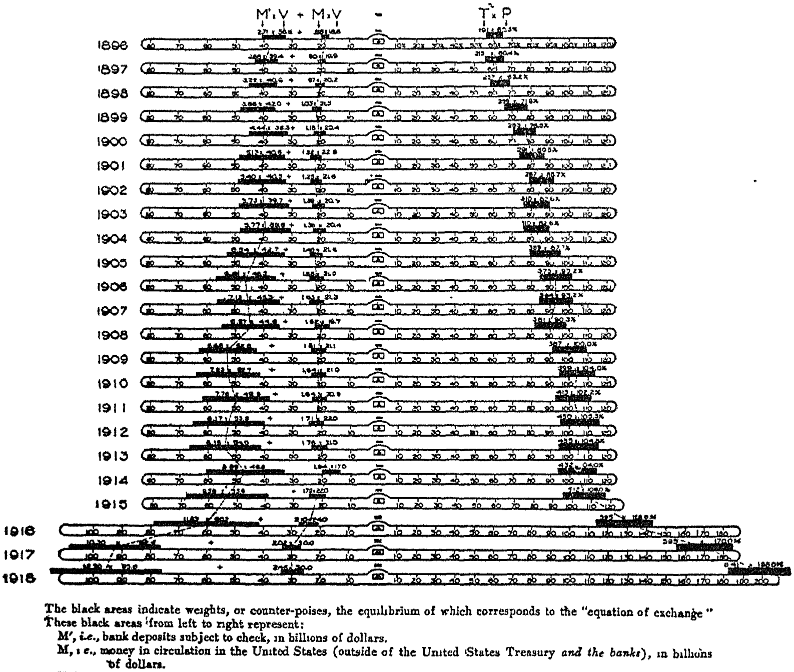
In both area-bar-charts and sky-line charts, it is sometimes useful to show by a dotted or broken line the contour of the entire group if all its varying areas were combined into one area. This light, broken line serves to show the average or normal or typical phenomenon from which the individual areas are variations.

When the items or stubs (the independent variable) of the data form an ordered numerical series, we find ourselves, in area-bar-charts, back to frequency curves in staircase form. The frequency curve in staircase form is essentially nothing more than a sky-line chart, generally of a certain characteristic type (the tallest area-bar being near the center and those at the side diminishing in height gradually until they vanish entirely). In the chapter on frequency curves we have seen that when the items form a continuous and not a discrete series, a smoothed curve can be plotted through the mid-points of the top ends of the area-bars making a frequency polygon. And we have seen that while the total area under the frequency curve truly represents the total aggregate of the frequency series, nevertheless the areas between any two ordinates under the smoothed curve are not individually equal to the corresponding area-bars in the staircase frequency curve (unless the adjoining values about any particular item happen to form an arithmetical series). This inaccuracy, we have seen, is sometimes more than compensated for by the more suggestive value of the smoothed curve.

A more complicated form of the area-bar-chart occurs when the areas are segmented like 100% bars to show a secondary classification or subdivision. This chart closely corresponds to the compound bar-chart, and the 100% square, differing from the former in the varying widths of its bars or layers, and from the latter in their varying lengths. It is generally of use in the analysis of the parts of a frequency series, either cumulated or simple. It is a sort of frequency band curve, in which the actual values are plotted on both axes, where the 100% square, rectangle, or marble-cake (smoothed vertical

curves) chart projected the secondary classification only in percentages. The "stream chart," too, in which the bars or areas are arranged on both sides of an axis, can be modified to present similar area features.

In the particular case where the data contains two pairs, with one product always equal to another product, the two areas representing these two pairs can be pictorially shown as suspended and balancing each other upon the two arms of a chemist's weighing scales or balance. The two areas must of course be suspended at equal distances along the lever arms of the balance. The pictorial representation of the balance suggests to the reader the equality of the two products.



The black areas indicate weights, or counter-poises, the equilibrium of which corresponds to the "equation of exchange" These black areas (from left to right represent:

M', *i.e.*, bank deposits subject to check, in billions of dollars.

M, *i.e.*, money in circulation in the United States (outside of the United States Treasury *and the banks*), in billions of dollars.

T, *i.e.*, the volume of trade circulated in billions of "units" (each "unit" being that quantity which could be purchased for one dollar in 1909).

The *lever arms* of the above three weights represent:

V', *i.e.*, the velocity of circulation ("activity") of the deposits, M'

V, *i.e.*, the velocity of circulation of the money, M.

P, *i.e.*, the index number, or scale of prices, at which the trade, T, is conducted (This scale of prices is measured as a percentage of the scale of prices of 1909.)

Fig. 476. Balance or Counter-poise Chart with Two Factors.

The chart illustrates Professor Fisher's quantity theory of money, according to which (M' x V')+(M x V) = (T x P) that is, checking deposits times their velocity, plus money in circulation times its velocity, equals prices times volume of trade. The weights are here shown as horizontal lines and their factors as leverages.—*Permission of Mr. Irving Fisher.*

A more striking method of using the same idea is to show one of the factors of each product by a weight or bar suspended from the lever arm and the other by the distance along the lever arm between the weight and the fulcrum or point of balance. In this case, we have, as it were, abbreviated the area, merely showing its two dimensions and leaving the reader to imagine the area itself by projecting these two linear dimensions. The picture has little analytical value but is a powerful means of visualizing to the reader the mathematical relation, $AB = XY$. Professor Irving Fisher has used a series of the two-factor scales or balances with excellent effect in his exposition of the quantity theory of money.

This method can indeed be extended to show the equality of products not of two, but of three, factors as well. Each pair of "balances" or weighing "scales" illustrates an equation of the form $ABC = DEF$. And a series of such charts would illustrate a series of such equations. Since we have used the radial distance from the fulcrum or point of balance to the point of suspension to represent one factor, we can show the weight there suspended as a two-dimension area, the length

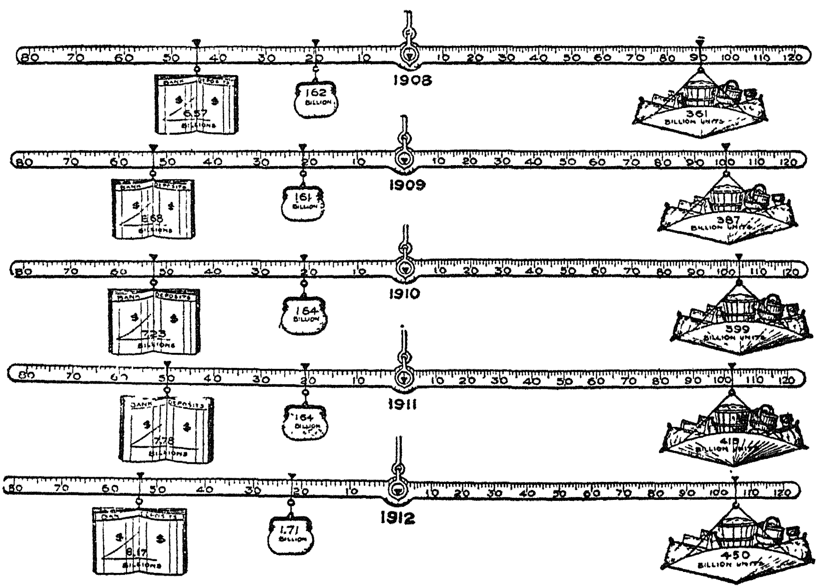


Fig. 477. A More Pictorial Form of the Preceding Chart.
A detail of the chart in Professor Fisher's book, "The Purchasing Power of Money," in which the second factors, that is, in the chart, the weights, have been pictured realistically.—*Permission of Mr. Irving Fisher.*

and width of which illustrate the other two factors. Such balance pictures need not be the same in pattern on both sides or areas of the balance; we may for example show the three factors on one side and their total product upon the other, using for the latter merely a horizontal bar. Thus total sales may be shown as a plain bar-chart, each bar being centered at a fixed point on one side of the balance, and per capita sales may be shown as an area-bar chart, positioned at the other end at distances corresponding to the population. The areas would show a long one dimension the "last year's per capita sales of the quota" and along the other the present percentage thereof. In short the balance or weighing scales chart is but a form of compound area-bar chart with especial pictorial value for its particular relations.

The ingenious chart-maker will be able to apply the principles of the area-bar chart in a wide variety of ways, always remembering that the widths and areas are less significant than the lengths and should be only used as qualifying or secondary information serving to evaluate or weigh the importance of the primary information shown by the length of the bars. In some cases, it is possible to apply these principles even to such well established bar-charts as the Gantt progress chart. A step in the direction of this qualifying evaluation was taken in the chapter on bar-charts when wider bars were recommended for total-group and sub-total bars in a bar-chart.

There is, indeed, no reason why the area chart principles cannot be applied to circular graphs. And in the next chapter the reader will find the same principles extended to wholly irregular areas, such as map outlines. In general, the area chart like the 100% square, is essentially popular in its appeal and simplicity is therefore of first importance. Although every digression from simple linear measurement results in a loss of precise legibility, yet when properly used, the principles of areas can be made to improve a great many charts both in attractiveness and in instructive value.

Chapter LII

POPULATION MAPS

Every digression from simple linear measurements results in a loss of precise legibility. In the rectilinear area chart we have seen that the area itself was but a poor illustration of the values it represented and was therefore useful chiefly for the sake of the general impression which it gave of relative importance of items or values already illustrated by lines. We are now about to take still another step away from simple linear dimensions and make use of areas of irregular outline. It is therefore more than ever necessary to repeat that the areas have little more than a qualifying or evaluating use, serving to give a general impression of the relative importance of items.

We do not sell our goods to the mountains, bill them to the rivers, or credit the forests with payment. Probably from at least a subconscious appreciation of this circumstance, many national distributors, advertisers, and sales-managers have discarded maps on which the rivers, forests or mountains are shown when they are studying the geographic distribution of their sales. The up-to-date sales manager plots his distributing points and records his sales in a great many ways upon maps which carry only faint State outlines or at the most show the location of the larger cities. But why stop here? Your sales manager does not sell to square miles, acres, or other units of land-area measurement. He sells to human beings. Why should he use maps which show, not human beings, but square miles, that is, maps in which the areas indicate not the population but the land surface? Why indeed!

The average density of the population in the United States proper at the last census was thirty-five persons per square mile. This density however varies from State to State. In some New England States there are more than four hundred

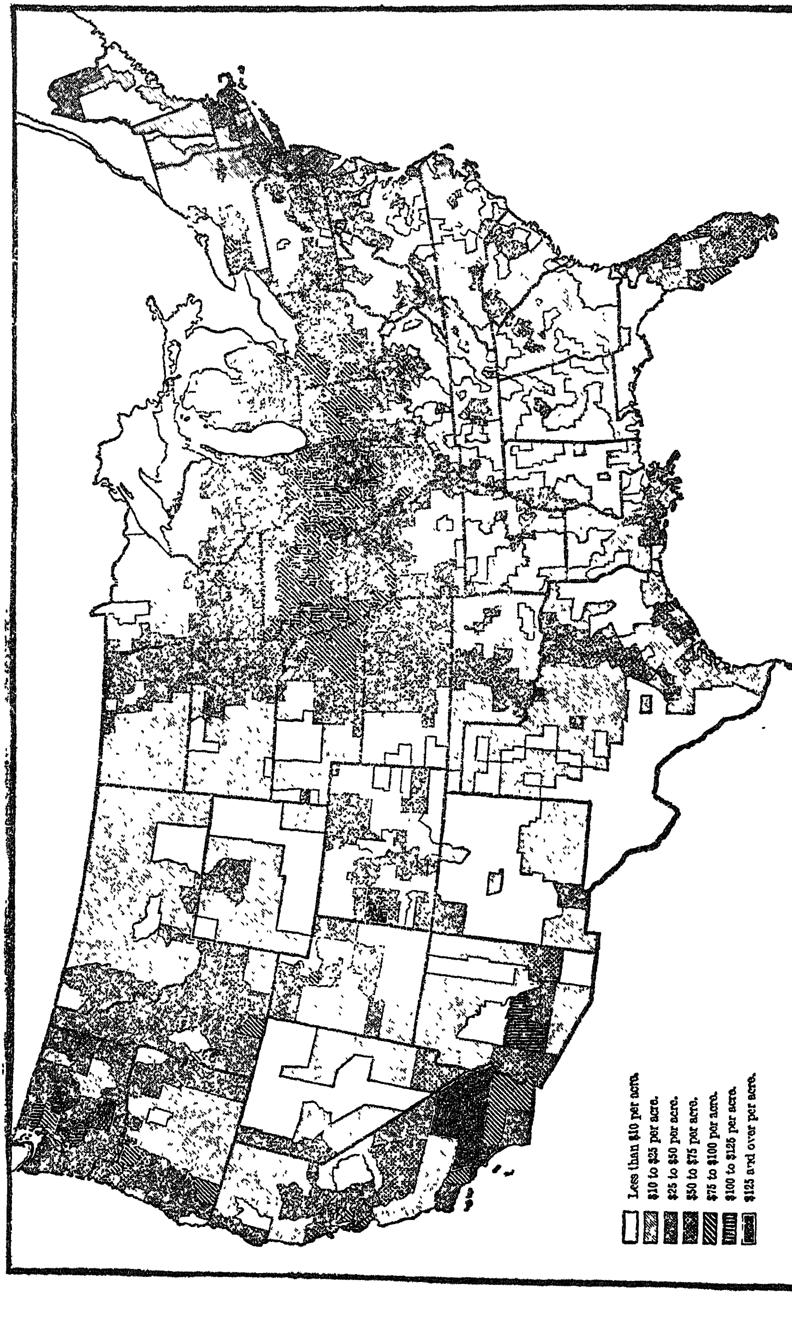


Fig. 478. Every Map is an Area Chart. On this Map the Areas Represent Square Miles.
Showing that the value of farm-land, shown by the shadings, is very different in different parts of the country.—*From U. S. Census.*

persons per square mile while in some Rocky Mountain States there is less than one person per square mile.

A handful of peas in the bottom of a box can be kept in a small corner if you hold them with your hands, but if you release them, they will quickly spread most evenly over the bottom of the box almost like water. Imagine the population which is pent up in these small eastern States, suddenly being released like the peas in the bottom of the box, and flowing out over the land of the United States until its density is uniform throughout, that is eight persons per square mile. Also imagine the population as carrying its State borders with it so that the enormous population of the Northeastern States, spreading out more than half way across the continent, would carry the borders of the northeastern States westward and southward with them. (Readers of this book who live west of the Mississippi river or south of the Mason and Dixon line may omit the remainder of this chapter!)

The result of this projection of the map of the United States upon a population basis rather than a land-area basis will be most surprising even to the most hardened travellers. A comparison of such a population-projection map with a land-area projection map will show how far the State lines have been shifted. From a position about one-third of the way across the map from the Atlantic ocean, the Mississippi river shifts to a position about a fifth of the way from the Pacific. The Rocky Mountain region becomes a narrow strip on the map. The Southern States shrink frightfully. But if the familiar outlines of the States are approximately kept in the new projection, the States will still be easily identified in their new form and you no longer have difficulty in locating important but crowded eastern cities.

Needless to say, the picture of sales conditions which such a map exhibits, will be far more valuable and useful than the picture upon the usual land-area basis. For in spite of a thorough knowledge of the various State populations, even an expert on population statistics will find less difficulty in visualizing sales conditions as far as the real market, that is the population itself, is concerned. You will no longer attach grave importance to the far Western States which show up poorly on your colored scale map, for they will no longer be enormous and terrifying red areas. But you will attach far more importance to the red color when it appears in the

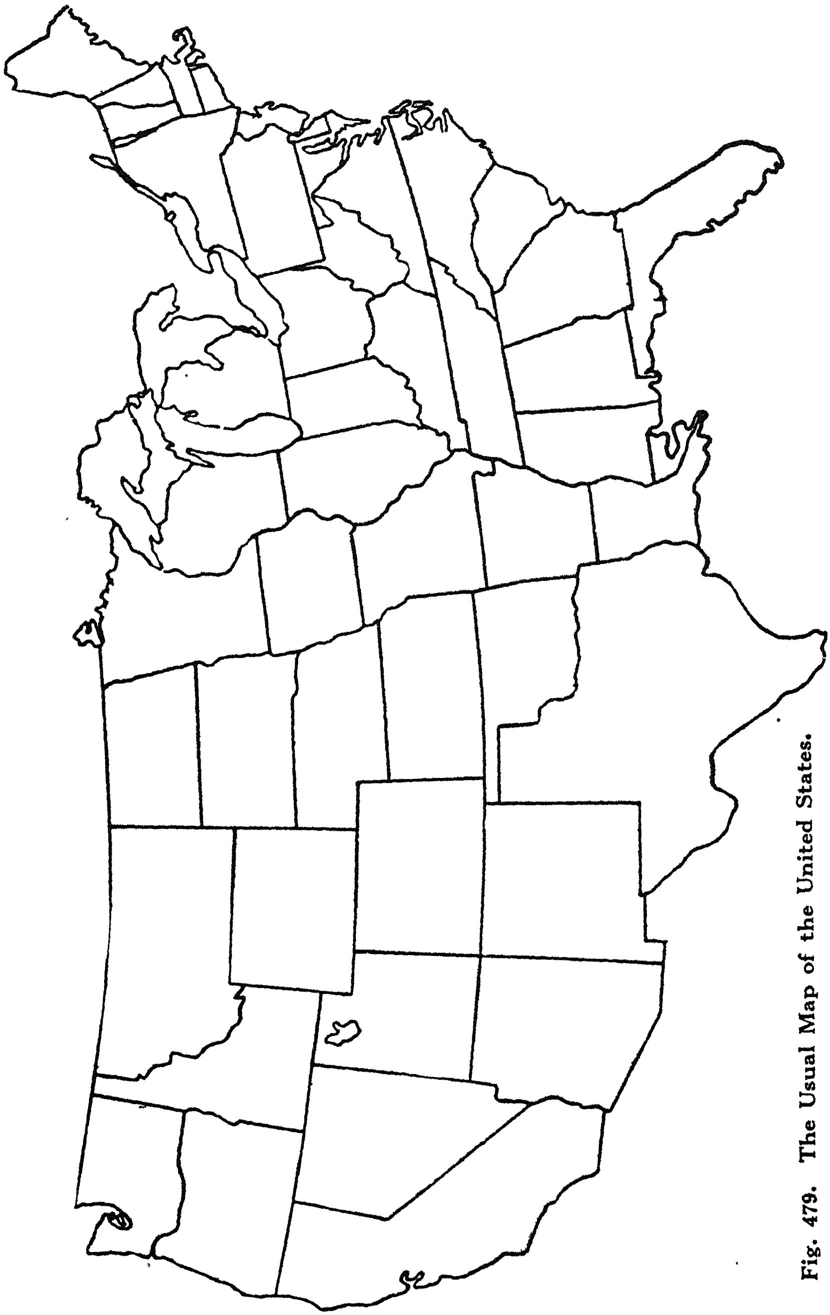


Fig. 479. The Usual Map of the United States.

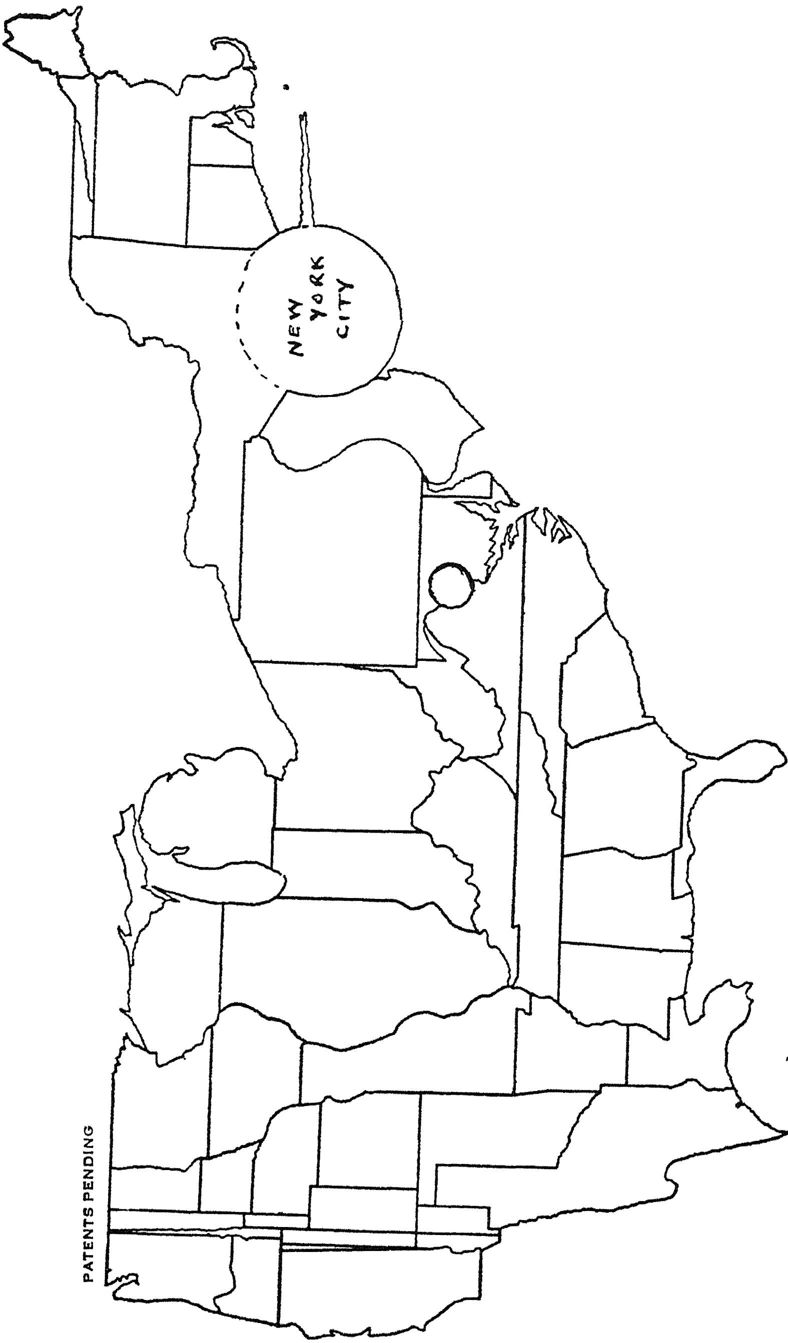


Fig. 480. On this Map the Areas Represent Inhabitants.

eastern States. Per capita sales statistics, especially, will be useful on this projection. And the location of sales distribution points, branch offices, and representatives or other salesmen will on this map show how evenly the market is saturated with your agencies rather than how regularly they are placed as mile-stones across the country. In short, the corrected areas of the States serve to give an excellent background or evaluation of the importance of the statistics plotted upon the map.

Other conditions beside the total population can be made as the basis of projection of the map. If your market is best shown by the native white population, for example, the map should be projected upon not a total population basis, but upon the basis of this particular class of the population. If your market is best shown by the dealers, retailers, jobbers or brokers, then the map should be based either upon the number of distributing agencies or upon their aggregate financial rating. For the analysis of business or economic conditions relative or comparable to the wheat crop, or any other form of produce or natural resources, a map might be called for, which is projected upon the basis of the average yield of this particular crop or resources during the past years.

The number of ways in which the map can be altered and projected for special purposes upon special bases is unlimited, but all are alike in one respect—that their areas no longer show physical land areas in square miles but show the actual values more important for the special purposes in view. This does not mean that a land area map is useless; on the contrary, there are many processes, such as shipping, railroading, and travelling, for which the actual land distances are important. It is only intended to show that where land area is not important and some other condition is important, it is possible for the map to represent the values we consider important, whatever they be.

The making of such special projection maps is very difficult and tedious and it is much better to purchase them from the few publishers who supply them. When you must prepare them yourself, perhaps the best method is to use a large sheet of cross-ruled paper in which the co-ordinates cut the paper into small squares of one-tenth of an inch or one millimeter each. Having before you a table of the values which you wish to project as areas on the map, lay out by the rule of "try, try,

try again," the State outlines, taking care to maintain their familiar shapes as far as possible, but at the same time counting the number of small squares included in the outline and changing the outline until the right number of small squares has been inscribed. For checking purposes, a planimeter (an engineer's instrument which measures areas) is useful. If only rough outlines are required, you might find that a large supply of differently colored small glass balls could be counted out for each State, one color for each State, and quickly adjusted into familiar outlines and closely packed to secure the right area. Another short cut is to use differently colored plastocine or children's modelling wax, and, having weighed the right quantities to suit your data, to mold these into familiar outlines and press them flat to a uniform thickness. When beads or wax are used, color distinctions should be maintained for the different States. Thin strips of paper along State borders may help to keep the colors from mixing.

The salesmanager will seize upon this map eagerly for sales analysis. The economist will often find it invaluable. Even the layman, with no charting or graphic analysis to make, will find it of absorbing interest. The correction which the map gives to our conceptions of State populations makes the map of real educational value, and the school geographies should have not only national but world maps upon such projections. The student will note that the principle of the map is the same as that of other area charts. He will recognize that precise estimates of the values represented by the areas are not possible, particularly as the shapes of the areas are irregular. He will see that in common with other area charts, the real value of the areas on the map does not lie in exact measurements of the values of secondary importance represented by the areas, but in the general weighting or evaluation of the relative importance which is given to the data of primary importance plotted or recorded upon the map.

Chapter LIII

MODELS

We have now seen used for charts, successively, the point, the line, and the area. The single straight line, or single system of straight lines, ending at specified points, forms the bar-chart in its many forms; if the line is circular, the pie-chart results; in all of these the points are perhaps the essential feature and the lines and areas may be considered incidental. A series of points or dots connected by a line, forms a curve; the plot is then of several points upon a dual system of straight lines, called co-ordinates; the outstanding feature here is a line (called a curve) to which intersect points and inscribed areas are incidental. A series of such lines or curves may be used to mark off areas, forming the area chart. But all of these forms are limited to the use of two dimensions. A third dimension is not supposed to be present, and is actually negligible, being no more than the thickness of the layer of ink, crayon, or color used in making the chart.

We now come to project points, lines and surfaces upward from the plane surface to get charts in which the third dimension itself is significant. And as may be imagined, we can attach importance and primary significance either to the elevated point, the elevated curve, or the elevated surface, to the vertical lines of elevation or the horizontal lines so elevated, or to one or another of the edgewise planes supporting the elevated plane, or to the entire volume itself. In short, we may use as significant any or all of the various intersecting lines and surfaces which make up the three-dimensional body, or the intersect points themselves, or the points, lines, and surfaces within the body. And there is something new—we can use the cubic content of the three-dimensional body as a basis of charting. This is a gain in simplicity at the cost of other things, and with this use of the three-dimensional body,

solid, volumetric chart, or model—call it what you will—we shall begin.

The reader whose mind has leapt ahead of the diverse forms of area charts to speculate upon the possibilities of three-dimension charts will not be surprised to find the "model" treated as a type of graph or chart. He will realize further that the model stands in the same relation to flat-charts as sculpture to pictures. Just as the floor surface of a room may be shown by the area of a plane representation having length and width corresponding to the length and width of the room, so too the cubic content of a room may be shown by the volume of a solid model having the length and width of the area chart and the further element of height corresponding to the height of the room. Just as the area-chart represented by its area, a product of *two* factors shown linearly, so too the solid model represents by its volume a product of *three* factors shown linearly. In both cases the scales for linear measures can be projected only arithmetically to secure the significant representation of the resultant or product, shown in square units in the area-chart; in cubic units in the solid model.

It is important to note that in the solid model even more than in the area, the representation of the resultant or product, though precise enough, is not easily amenable to precise estimation or comparison with other such products, for the human eye can even less easily judge of the relative values of two volumes than it can of two areas. Hence volumetric measures should be used, as square or surface measures, only for data in which the products themselves are of less importance to us than the factors which we will show linearly and which go to make up these products.

The reader who has caught the relation between the curve-chart and the area-chart in the realm of two-dimension charts will be prepared for a similar distinction in three-dimension charts. Comparable to the curved-line chart would be the curved-surface chart; analogous to the area with square units of measurement, would be the solid with cubic units of measurement. Such a distinction, however, is not of great importance; in neither two or three-dimension charts is it a hard and fast division. In the discussion of area charts, we have already seen that whenever the areas are classified by, arranged in, or dependent upon an ordered numerical series, the areas may be fitted together to form a curve or curves. Like-

wise in the three-dimensional charts, whenever the solids are classified by, arranged in, or dependent upon an ordered numerical series, we can fit all the solids together to form a curved plane. But because we shall not consider as a special subject the single isolated cubes or solids, we shall have little use for a distinction between volumetric charts (in which the unit primarily is cubic measurement); they could be made, but rarely with profit. The more complicated structure of cubes, their more difficult presentation and inspection, and the fact that their outer surfaces hide their inner transverse planes which are essential parts of them, make such isolated solids and even sets of isolated solids, of little practical value. Under this head we shall therefore discard all consideration of segmented cubes. The experimentally minded will be able to construct not only 100% cubes but even sets of several cubes and solids of various dimensions, which latter he will be able to compare by means of either or all of their three linear dimensions, their three areas in square measurement, or their one volume in cubic measurement. For the comparison of different buildings, engines, machines or other physical equipment or structures, such models, or small replicas may indeed be useful. But apart from miniature replicas of actual physical objects (in three dimensions) there would be little use for such isolated models or sets of isolated models. For mathematical statistics, the individual factors are better compared in separate sets, in bar-charts or curves, and the products are likewise better charted by linear measures in bar-charts or curves.

Just as the area, measured in square units, is not significant in all two-dimension charts, so too, the volume, or cubic content is not always significant in three-dimension charts. For the co-ordinates in a curve-chart, for example, need not have any factorial inter-relation; this is the case of historical curves; we do not multiply the phenomenon by its date of occurrence to secure a significant product. Likewise the three systems of co-ordinates used in projecting a solid, need have no factorial inter-relation which has a meaning for us; in such cases the three axes and sets of co-ordinates are merely convenient plotting devices which enable us to distinguish three different variables in our data, and to study their mutual relations and behavior. At other times, we may detect a distinctly factorial relation between these variables, and then, of course,

we can identify the product with the cubic content or volume of the chart.

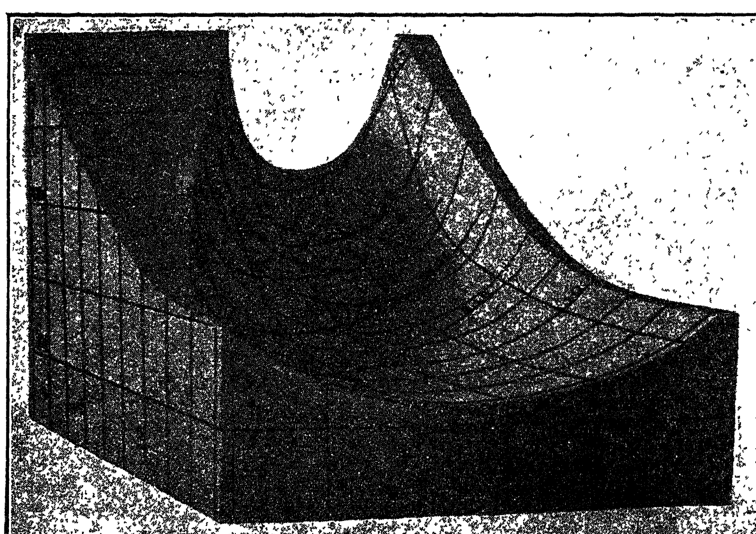
The most convenient division of three-dimension charts, however, and the one which we shall here follow, is the division between three-dimension charts in general, and one particular kind of three-dimension chart in which the first two dimensions are used to mark off geographical relations. To all other models and three-dimension charts we give the name of frequency surfaces; in these the first two dimensions have for their scales, any other numerical series, either historical or frequency. But when the numerical series indicates latitude and longitude upon the earth's surface, or any geographical co-ordinates, we call the chart a map and because of great practical use which is made of maps, we shall consider them in a separate chapter. As has already been pointed out, the isolated solid or series of separate solids, which is comparable to the bar-chart when the latter cannot be converted into a curve, that is to say, which is not classified by, arranged in, or made dependent upon an ordered numerical series and so cannot be joined into a curved surface, will not be discussed at all. And before proceeding to the consideration of the two types of three-dimension charts we shall first examine in the next chapter, the methods by which the third dimension can be shown.

Chapter LIV

THE THIRD DIMENSION

There are three ways of preparing stereographs, that is, three-dimensional charts, which may be called respectively, the model, the axonometric chart, and the orthographic chart. The first requires three dimensions physically in space; the second illustrates three dimensions in a two-dimension plane; the third shows two dimensions faithfully and seeks to represent the third dimension by some trick of symbols.

If we elect to use the model, it may be either solid or collapsible. To the solid model, actually built up in three dimensions, there is of course little structural difficulty. A solid



From "The Construction of Graphical Charts," by John B. Peddle, published by McGraw-Hill Book Co., N. Y.

Fig. 481. A Plaster of Paris Model.

Model showing the relation between heat units per hour per brake horse-power, compression pressure, and volume of gas mixture for a gas engine.

model can be made of wood, or of many layers of cardboard or corrugated paper, and can also be moulded of plaster of Paris, of wax, paper-pulp, or other material. Detailed instructions for the modelling of curved planes will be found in the following chapter. Such solid models are of course cumbersome and unhandy, difficult to file away, or to carry about from place to place; and would seem justified only in the case of extremely important data. Moreover, and this is important, the making of such solid models requires a great deal of time and trouble, the equipment not being generally immediately available for them in the average statistical office.

Special varieties of the solid model will also be described in the immediately following chapters.[1] These consist of small forests of vertical wires, or wooden sticks, placed far enough apart to allow any individual wire or stick to be inspected. This is in some respects the model par excellence; it admits of segmentation, for each wire or stick can be differently colored through parts of its own length. And most unusual of all, several models can be combined by placing their wires or sticks side by side with distinctive coloration. The coloration of wires is usually achieved by stringing colored beads upon them. The forest model is a more or less laborious affair to construct, but for sufficiently important facts it is of ample merit to justify its use.

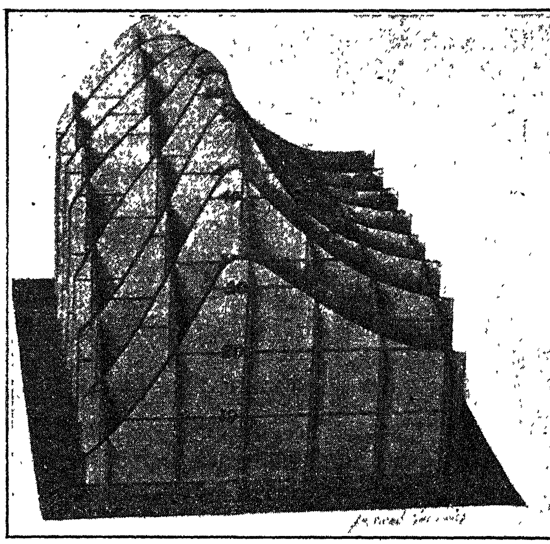
Collapsible models, as the name suggests, are so made that they may be folded up or expanded at will. When folded, they lie flat upon a single sheet of paper and can be easily carried about or filed away as sheets or folders. When expanded to occupy three physical dimensions in space, they can be stood up like blocks or other solids. There are three types of collapsible models, with respect to the mechanics of their operation. The first type telescopes by means of co-planar hinges, at right angles to each other, like the folds of a pair of bellows or the hood of a folding camera. The second type collapses side-wise upon hinges which are all parallel to each other, like the partitions in a pasteboard egg-box. The third type never opens out fully for all parts of the model are hinged together like the leaves in a book.

---

[1] See also, Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Magazine Co., New York. Peddle, John B., *Construction of Graphical Charts*, McGraw-Hill Book Co.

In the telescopic model all three dimensions are physically represented by materials in the structure of the chart and telescoping is only possible by buckling up of these materials across one of the three dimensions. Because of this buckling of materials the telescopic model is not easily made or operated, and is generally inferior to the other collapsible models.

In the second type, collapsing side-wise, because one of the physical dimensions is not represented by any structural material in the chart, the materials lying on the other two physical planes can be made to fall together as easily as a



From "The Construction of Graphical Charts," by John B. Peddle, by permission.

Fig. 482. Collapsible Model.

house of cards. If the hinges are on edge the chart folds out to right or left. In this case it is most convenient to make the chart of intersecting slitted sheets, the sheets to stand in one direction having vertical slits or key-ways, halfway up from the bottoms at the points where the cross-wise sheets intersect, the latter sheets having corresponding slits down through their upper halves; and the two sets fitting together, as has been said, like the partitions of the egg-box in which you buy a dozen eggs. Obviously all sheets should be of stiff material, rigid enough not to fall over, strong enough not to tear beyond the slits.

When this type is made with horizontal hinges, only one set of vertical sheets is used and these are all hinged, parallel to each other, upon a single horizontal sheet of pasteboard. That the various hinged sheets may act as one and maintain their parallelism at all times, short tie-hinges or keys and key-ways are used across or near their tops. It is convenient to mount this chart in a heavy pressboard folder of the standard vertical filing type, with the back of the chart attached to one half of the folder and its base to the other so that the chart is always safely housed. Such a model opens out like the more elaborate, familiar valentines and works on the same principle. Similarly the forest model already mentioned can be made collapsible, paper being used in the place of wires or wooden sticks. This type is perhaps the best of the precise collapsible models.

The third type of collapsible model has no structural materials for two of the three perpendicular planes in the three dimensions, and, as has been said, never opens out fully. Its single system of plane surfaces, which should be parallel, radiates from a common hinge like the leaves of a book. It is in fact no more than a series of two-dimension charts carefully bound together to secure perfect "registration," one with another. But frequently it is sufficient for the study and analysis of the data; and the fact that it is more easily made, and handled, and suffers less from wear and tear, strongly recommends it. Moreover, transparent or semi-transparent paper can safely be used for this chart, enabling the reader to note more easily changes in its various parts.

This last form of collapsible chart is also readily susceptible to commercial publication. In German schools and colleges such diagrams are sometimes used for the study of parts of complicated machinery. The student no longer needs to have a physical model of the machine before him, but can fold back the diagram of each part to inspect the diagram of the parts inside it. In medical schools such methods are sometimes used for the illustration of anatomical studies, the first view showing the outer skin; the second, the underlying nerves; the third, the muscles; the fourth, the inner organs, with perhaps minor diagrams or part pages folding back for each of these, to show the internal structure of these organs; the fifth large sheet (seen by folding back the fourth sheet which carried these minor books or sets of small pages upon it) showing the bones;

the sixth, the rear wall of the muscles, nerves or skin again. The same method of presentation has been effectively used in costume studies for the stage (and in children's toys), to show upon a top sheet the over-garments and outdoor costume; upon second sheets, the ordinary or house garments; and upon third sheets, the undergarments, for various national or period costumes. There would seem to be no reason why the same method cannot be used to present mathematical data in similarly constructed and hinged charts. For these books or series of superimposed charts, as for the collapsible models, a strong cover is desirable to protect the parts and the best cover is generally, as has been said, a vertical filing pressboard folder.

All of these model-charts have a bulkiness, when solid, with an added element of flimsiness, when collapsible, that militates against their general usefulness. Moreover, in their preparation they are costly of time, and often seem to call for material, not readily available in the average office. Except in the last form, when transparent paper can be used, or in other forms when transparent celluloid is used, or in the solid model when glass is used, these charts have the disadvantage that one part of the chart hides other parts; and diagonal lines, curves, or planes cannot be readily run through the chart to give interpolated readings. These disadvantages all disappear in the next form of three dimensional chart, which we shall now consider.

The axonometric chart is one in which distances are measured along three axes which have been represented by lines within a single plane surface. Every photograph and every picture of physical objects is such a chart. In paintings, drawings, and pictures of all kinds, perspective is generally introduced, to make the more distant objects smaller, that they may appear to be of the same size. Perspective requires that really parallel lines be shown as converging toward a non-existant vanishing point; it makes a fixed scale for any axis useless and greatly increases the difficulties of drawing. For statistical charts, therefore, we omit perspective, sacrificing thereby some of the realistic appearance of our picture, but giving it constant scales which make charting easy and reading accurate.

The most commonly-used axonometric chart is the one with isometric rulings.[2] Isometric rulings are those in which the three axes of the chart intersect to form equal angles of 60 degrees with each other. Any other angles can be used but these isometric angles are most convenient and satisfactory, as they afford the fullest detail along each axis with the sharpest possible intersections of all co-ordinates. One of the axes is vertical for the up and down dimension of the chart; the other two, at 60° to this or 30° to the horizontal, represent the two surface dimensions of the base of the chart.

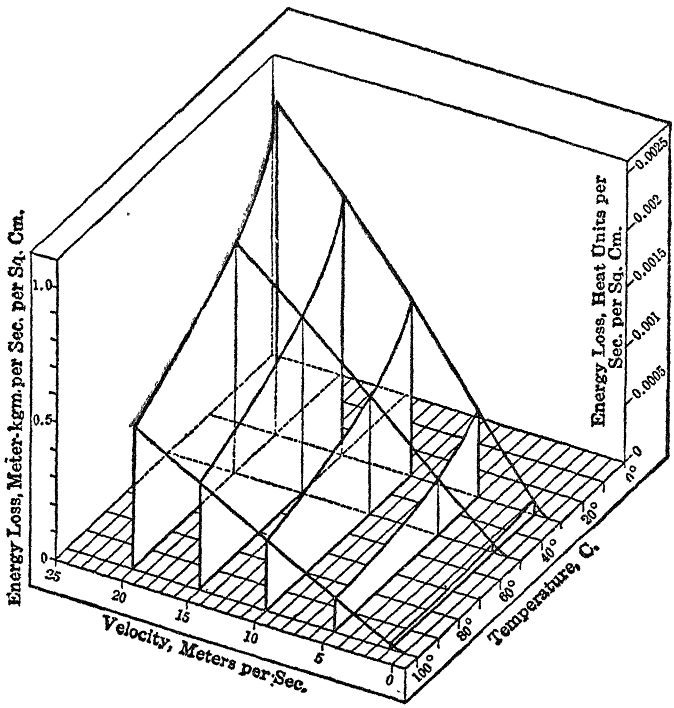
Isometric drawings are so very useful and so easily made that they should be adopted whenever possible for three-dimension charts. The isometric co-ordinates can be left upon the finished charts to facilitate interpolation and estimates of the values plotted linearly along them by the reader of the chart, for the rulings serve to connect the points plotted, with their scales, in the same way that the Cartesian co-ordinates, that is, the ordinates and abscissae, connect points upon a curve with the scales of the curve-chart.

But often the effect of many co-ordinate rulings is so confusing in the finished chart that it is preferable to wipe out all the co-ordinates save those which have been calibrated with scales and those along which curves or lines have been plotted. When the isometric co-ordinates are to be omitted in this way from the final chart, it is better to rule the co-ordinates in the first place on blank paper in pencil, as they can then be most easily removed. The commercially-ruled paper is printed in ink, generally green or red, and the lines will be reproduced in photographs unless wiped out with Chinese white. When the drawings are to be traced, of course it is very easy to omit the co-ordinates. But isometric co-ordinates are so easily prepared in pencil that for most purposes this is sufficient.

The scales for the isometric chart axes can be varied, of course, at will; and the student will naturally seek to adjust them somewhat to the ranges of the variables plotted. But when all dimensions, or even the two surface dimensions, are used to present commensurable values, and the commensurable nature of these is manifest, the use of different scale-moduli or units of distance along the different axes, may result in an awkward unnatural appearance to the chart

[2] See also, Professor Guido Marx in the *American Machinist* Vol. 31, Part 2, p. 701, and Haskell, Allan C., *How to Make and Use Graphic Charts.*

unless the angles of the axes are altered and the isometric principle departed from. It then becomes advisable to select other angles, which restore the natural appearance by swinging the whole chart about to one side or the other. The first consideration is, of course, the range of the variables; the second, the relative detail with which the variations should be shown. From these the total length of the chart along either axis can be determined and from the relative detail alone, that is the size of the unit distance or modulus, the proper angles of the axes should be determined. When it is desired to vary the angles for these purposes, the ready-made isometric paper of course cannot be used and the co-ordinates must all be especially drawn.



From "The Construction of Graphical Charts," by John B. Peddle.

Fig. 483. An Axonometric Chart (Not Isometric).

Chart showing relation between journal-bearing temperature, surface velocity, and heat generated, etc.

It is a distinct limitation of all drawings of solid objects, including isometric drawings, that they show the object from

Ratio of scale-moduli or units of length along the three axes.			Tangents of angles formed by the right and left-hand axes with the vertical axis.	
Left-hand axis	Vertical axis	Right-hand axis	Left-hand axis	Right-hand axis
mz	my	mx	(z)	(x)
1	1	1	tan 60°	tan 60°
			(Isome	tric ruling)
2	1	2	8:1	8:7
3	1	3	18:1	18:17
4	1	4	32:1	32:31
5	4	6	5:1	3:1
9	5	10	11:1	25:8

From John B. Peddle, "Construction of Graphical Charts"

Fig. 484. Instructions for Axonometric Chart Scales.

one view-point only. You cannot turn the drawing over for different views of the same object, as you could turn the actual object itself about in your hand. It is therefore necessary to arrange the scales of the axonometric drawing carefully to show the best possible view of the object. High points in the foreground will hide or obscure lower points behind them. The scales should therefore be so arranged that the high points will be set as much as possible in the background, and the foreground be devoted to low points; or that sufficient distance be allowed behind a peak in the foreground to enable us to show low points behind it. If, by reversing the direction of one scale, you can secure this result, the scale should be reversed; for the result of reversing a diagonal scale is the same as giving the object itself a quarter-turn in your hand. The reversal of the other diagonal scale is the same as giving the object itself a quarter-turn in the opposite direction; and the reversal of both axes gives a half-turn to the object itself, showing its rear face. In general, the best position can be found by a little experimenting and can, with a little skill, be determined in advance from an inspection of the data.

It is, however, an advantage of the axonometric chart, not shared either by the model or the orthograph, that interpolation can be most easily accomplished upon it. For the axonometric chart can be made, if we wish, to show all sides of the chart at once, merely by using points or dotted lines for the parts which are apparently hidden from view. And even if these parts are not indicated, the scales and axes still remain, from which we can on the finished drawing drop parallels to any desired point and take a reading. Such interpolation

can be taken from straight lines in the manner just mentioned or from rounded contour lines drawn in from various observations, according to the nature of the problem. The axonometric chart, and in particular the isometric one, is for most purposes the most satisfactory method of charting three dimensions.

A feature which belongs both to models and axonometric charts is that they may present either staircase or smoothed surfaces. This is obvious enough from the fact that both are but series of curves; and curves, as you know, can be in either form. Moreover, since the curved surface or three-dimension chart is an interlocking of two such series of curves, one along and the other across the surface, it is obvious that the same surface may even be smoothed along one axis and staircased across the other, as well as smoothed or staircased on both. These possibilities are not generally open to the charts to which we are coming, the third type of three-dimensional chart; in the latter the best that can ordinarily be done toward smoothing is to indicate contour lines, or lines of equal value, about each peak and valley—a lateral, if you will, rather than vertical method of smoothing.

The third method of presenting three-dimension charts (they can hardly be called stereographs in this case) is the orthographic chart. In this, as already mentioned, two of the dimensions are precisely shown, exactly as in a two-dimension chart, and the third dimension is indicated by symbols. In its very simplest form, numbers alone represent the third dimension, the numbers being for this purpose considered as symbols. But the graphic quality of a number is limited,—consisting wholly of the number of digits in the number itself,—and this is not usually sufficiently detailed or legible. We are therefore prone to seek other symbols to which we can attach numerical values and which we can explain to the reader of the chart in an appended "key" which corresponds to the scale along an axis. And since no symbols have yet been found which are capable of the infinitessimal graduations which a scale affords and are at the same time as easily read with accuracy, we are obliged to restrict ourselves to a few distinctive symbols. These we use not only for certain set values, but also for all values nearest thereto, establishing in this way groups or intervals along the range of variation and using these symbols for all values within each group. In short,

the use of symbols involves a loss of detail in the presentation of the third dimension.

Two considerations govern the use of the symbols. The first is that the groups to which the symbols are attached should be carefully chosen. This consideration is precisely the same as applied to the formation of frequency series.[3] The groups should if possible contain any round numbers or bunching-up spots near their centers. The intervals, or group-limits, should be regular and uniform, if the distribution appears arithmetical, and as nearly as possible to equal geometric intervals if the distribution appears to be logarithmic. These are obvious principles which the student will soon discover for himself. The one thing of consequence is that we should not, as we may often be tempted to do, divide the series into groups with equal frequencies. Such a practise is confusing and deceiving to the chart-reader and has but limited meaning.

The other consideration is that the symbols should be such as form a natural series in themselves, just as if they were numbered. This gradation of the scale of symbols should be such that it is obvious to the chart-reader—the more obvious we make it, the better is our representation of the missing third dimension. The reader of the chart should be able to see at a glance the order in which the symbols fall and thereafter should not need to refer to the key again. Indeed, in those few portions of the chart where the extreme symbols are used, it is no bad plan to label the chart itself, right through the symbol, with the words "High" or "Low," or with similar words. When this is done, the chart may be called self-contained and complete and the reader need only refer to the key for the numerical equivalents. Needless to say, however, the key should always be attached to the chart, showing the symbols and stating the limits of the ranges they represent.

The position of the symbols upon the chart is of course dictated by the two independent variables in the data, the plot of which occupies the two co-ordinate dimensions of the chart. But since the dependent variable can only be shown approximately by symbols, not precisely, there is opportunity for two different methods in marking off the parts of the chart

---

[3] See Chapter XXVII, pages 312-313.

to be symbolized. In the first place, we may take these parts precisely as we find them in the data. But such a method leaves to the chance boundaries of the data the question of what approximate values shall be shown by symbols for any particular spot. And as a result the most abrupt transitions of values may take place between two parts of the chart appearing side by side. Where each part is inherently homogeneous throughout, the resulting chart, though much confused, is nevertheless accurate.

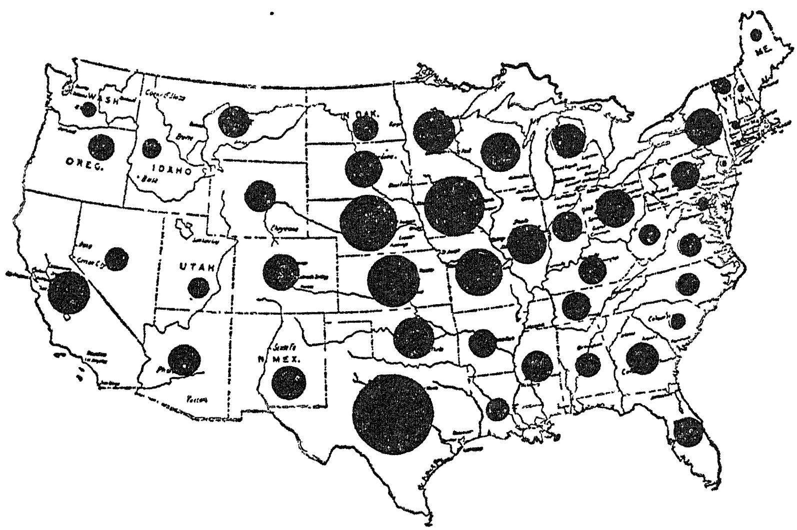
But where we have reason to believe that the change from one spot on the chart to another is more or less gradual, we are justified in trying to smooth out the steps between parts of the chart, so that all intermediate symbols appear between any two non-successive ones; in other words, so that the change from one part of the chart to another is as smooth as the few symbols and the given data will allow us to make it. This results in a much less confusing picture, and in the circumstances prescribed, a more significant one. It is the old distinction between a staircase and a smoothed curve, with all the attendant details of loss of absolute accuracy over given areas and greater significance. Only in the three-dimension chart of the kind we are considering, the process is called "zoning" and the lines which bound the zones are called contour lines.[4]

This is the orthographic chart proper, familiar enough in weather topography, where the lines of equal barometric pressure are called isobars; and the lines of equal temperature, isotherms. But to apply those principles to the chart, whatever it be, is often a difficult task. When data is so scattered that many contour lines must be interpolated between two known points, the element of "guess" becomes large, different chart-makers will often connect zones differently and it becomes often a decidedly hazardous proceeding. In such cases the method need not be employed, or if used, the interpolated zones may be made disproportionately narrow to limit the possible error as much as possible.

Many and various are the kinds of symbols which may be used. One of these is so close to the mere number itself, which we have already mentioned, that it may be disposed of first. This method consists of the use of bars, areas or circles in the

[4] Contour-lines may be considered a form of superimposed cross-sections.

place of the numbers. These all have the possibility of infinite gradation, just as have numbers themselves, and so form an exception to the considerations just laid down for symbols, and require no key. They are not wholly satisfactory, however, for they involve as essential the use of one or both of the dimensions already given on the chart to other variables. Hence the symbol cannot be evenly spread over the entire part of the chart to which it applies and when parts are small and symbols large, the symbol will cover parts to which it does not belong when perhaps on the same chart, other parts are large and have very small symbols, the symbol is likely to be lost and is not fully graphic.

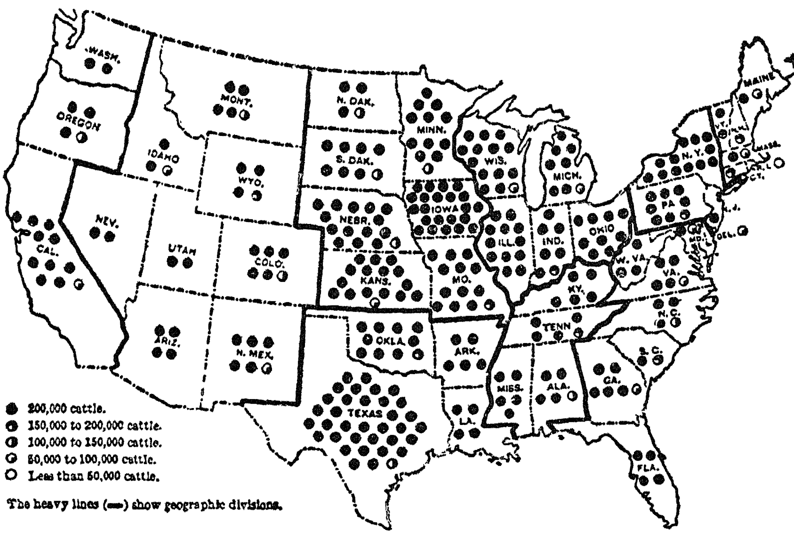


Permission of Country Gentleman.

Fig. 485. The Wrong Way.

If it is desired to use such symbols as these graphs-within-graphs, then surely the symbol should not be measured in areas, such as the squares, triangles, stars, or circles, which one so often sees used for these purposes. For the comparison between a large circle and a small one, or a large star and a small one, cannot be accurately made by the reader. It is much better to use bars or segments of circles, all requiring linear measurement only. For in this case the reader can be trusted to arrive at approximately accurate conclusions, in spite of the fact that the bars have not been aligned at one end. The one case in which areas (in square units) should be

used seems to be the case in which the variations symbolized seem to increase, not in arithmetical or geometrical series, but in a series of squares and the special square-root projection is desired for the scale of the symbolized function or third dimension.



From U. S. Census.

Fig. 486. Somewhat Better.

A logical development of this is the use of many dots or small circles in the place of one big one for each symbol. By counting dots the reader can get the exact value of the variable plotted in each part of the chart, unless the dots are allowed to become so numerous that they cannot be counted. This last trick, of putting in a great many dots for a single symbol, is unfortunately a popular, though pernicious, practise—it is as if the chart-maker were saying to the chart-reader, "I have gone to a great deal of useless labor in putting in all these dots, now you can waste your time counting them."

When the parts of the chart are of uniform size and the dots are evenly distributed within each part, the dot system is excellent; for the density of the dots is a graphic guide in itself; but when the parts to be labelled with symbols are not of even size, and no significant relation holds between the size of the parts and their symbols, then unhappily the dot-system falls down again: for a few dots in a small part of the chart will be more impressive than many in a large part.

The poly-dot symbol leads us logically to the frank use of shadings, regardless of the number of dots, lines, or other markings in a shading. And this is ordinarily the most satisfactory of all symbols. For a few different kinds of shadings can be easily devised, which are not only mutually distinct, but also have a definite order of intensity, ranging from solid white to solid black. These shadings can be laid on with successive hatchings and cross-hatchings, with a section-liner or tee-square. They require least work if the shadings are so chosen that each successive symbol has only an added system of lines or other markings to distinguish it from the previous one, for in this case all of the lesser shadings can be put on in the course of making the extreme shading.

It is an advantage of the cross-hatching symbols that they do not, as a rule, interfere with lettering which may be also wanted on the various parts of the chart; the lettering can be read through all but the solid black or very dark symbols, and in the latter cases, the symbol can be omitted immediately about the lettering. It is also an advantage of the hatched symbols that they can be reproduced, in common with the methods already described, in a variety of ways, including the line-cut for printing and the mimeograph for offsetting, and the blue-print or Van Dyke print for copying.

Closely akin to the hatched symbol, is the solid shade or tint, the shading proper ranging from pure white through various grays (made by mixing those two ever-present visitors, India ink and Chinese white) to solid black. These tints could, of course, be infinitesimally graduated to suit precisely the values plotted, but this would require unnecessary work and could not, through optical illusions, be correctly read by the chart-reader. It is therefore sufficient to use some five or six equally different shades which can be easily distinguished on a key. The method of solid shades is not, however, generally of enough benefit to warrant its use. It requires some labor in mixing, the liquid may warp the paper or run, the symbols are never so distinguishable as hatched patterns, and the resulting chart cannot be reproduced by line-cut, or any other method except photo-engraving or photostat.[5]

The most important and satisfactory symbols possible are solid colors. These can be made very pale so that lettering

[5] The "Ben Day process" can be used on a line-cut to make it slightly resemble a half-tone. See Appendix C.

or even a separate scheme of cross-hatching can show through them. Transparent inks and water colors, used in color photographs, can be used. Better than ink or water colors, however, are wax crayons, the cheaper and waxier the better. They do not require careful mixing, lay on in even density for each color and do not wrinkle the paper. After the wax color has been heavily applied, the chart should be carefully scraped with a sharp knife (a safety razor blade is excellent for the purpose) and all the surplus wax removed. A pale tint will remain on the paper, through which typewriting or other letters or chart drawings (if previously applied) will show clearly. The important thing about the scale for colors is that the colors should be in what is called chromatic sequence, the order of the colors in that of a rainbow or spectrum. Using five colors, red, orange, yellow, yellow-green and blue-green will be found excellent. For more colors a dark red and a blue can be added. But five clusters or symbols are ordinarily sufficient, yellow representing "average or normal"; orange, "poor"; red, bad"; yellow-green "fair"; and blue-green, "good."

Some writers have advised an arrangement of colors by what they call optical density and have attempted to determine a color density sequence. These efforts have naturally and necessarily failed—such a scheme, even if it could possibly be standardized for different inks, papers, and color combinations, would only result in conglomeration through which the reader would need the constant assistance of the key. The only disadvantage to colors is their varying photographic reproducing powers, blue disappearing wholly and turning white, while red becomes black. A careful chromatic scale through the colors from red to blue will photograph as a fairly uniform scale of grays from white to black. But the chief advantage of the chromatic arrangement of the colors is their logical significance. The reader of the chart if he be not color-blind, need only know that blue is good and red is bad and is at once prepared to interpret all the intermediate colors.

You will see that colors, shadings, and even figures alone, constitute a dimension in themselves upon the chart. And in almost all instances where models are made for three-dimension statistics, the same could be charted upon a two-dimension chart with the use of colors, tints and cross-hatchings in the place of the third dimension. That the symbolical presentation of the third dimension is more a series of approximations,

no precise graduations being possible, has already been explained. But in most cases these approximations are entirely sufficient.

When a more graphic or vivid representation of the same three-dimensional data is desired, with perhaps more precise presentation of the exact values of the third dimension, the isometric or other axonometric drawing must be used. But in this a part of the data may be hidden behind peaks. If this is the case and all parts of the data must be visible, then of course you must fall back upon the three-dimension model itself, either in collapsible or rigid form. The great time consumed in the making of these and the inconvenience in handling them makes the three-dimension model, rigid or collapsible, justified only in the case of very important statistics, but the ease and convenience of colors and isometric projections make them of very wide general usefulness.

Chapter LV

FREQUENCY SURFACES

The double frequency series is a type of data which can invariably be recognized by the form of its tabulation. It is composed of several columns of figures which have common stubs, and in which the values of the stubs and the values of the column headings both form mathematical variables. In the body of the table, that is, at the intersections of rows and columns, appear the values of the functions or, in a loose sense, the dependent variable. The general form is the same as the table of original data for the 100% square, already discussed,

WET AND DRY MONTHS

Summary of the number of times each month has been first, second, third, etc, in order of humidity during the years 1868 to 1905 — 38 years in all. Taken from flow of Croton River, N.Y. at dam.

	Jan.	Feb.	Mar.	Apr.	May.	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
Wettest	5	8	13	7	1						2	2
Second	1	7	12	6	2			2	1	1	3	3
Third	8	8	4	5	1			1	1	1	2	7
Fourth	8	3	5	8	3			1	1	1	1	7
Fifth	4	4	2	6	9	1		2	2	3	1	4
Sixth	4	3	1	1	7	2	4	1	1	4	6	4
Seventh	4	2		2	5	2		1	2	2	11	7
Eighth	1	1	1	2	6	11	1	2	4	5	3	1
Ninth				1	1	11	6	8	1	6	3	1
Tenth	2				1	5	8	4	5	7	4	2
Eleventh	1					4	9	9	9	4	2	
Dryest		2			2	2	10	7	11	4		

Fig. 487.

but the data is not turned into percentages or products of percentages as in that case. It is only necessary that the stub

and column heading figures, that is, the two independent variables, each form ordered numerical series. Whenever this is the case, the values of the function (that is, the detail figures of the double frequency series, shown in the body of the table) can be charted in a third dimension. If you think of pins stuck into the table upon each figure in the body of the table, the pins representing the figures by their heights, you will see at once how this is done.

Let us assume that we are standing in a room of rectangular or square floor shape. Let us mark off a pattern upon the floor of this room, a pattern of criss-cross or co-ordinate lines, calling those which run the length of the room the x-abscissae and those which run across the room the y-abscissae. At the many intersections of these two sets of abscissae, let us drive tacks into the floor and into the ceiling overhead and run strings vertically from the floor to the ceiling. Let us call these vertical lines the z-ordinates. At equal distances up these ordinates, or vertical strings, let us fasten horizontal strings from ordinate to ordinate above both the x and the y abscissae so as to produce a complete net work of crossed lines in the room, which would present the appearance of plain co-ordinate rulings when seen from above or from either side. Let us assume that in some mysterious way we can wander about this room freely without becoming entangled in the net work of strings.

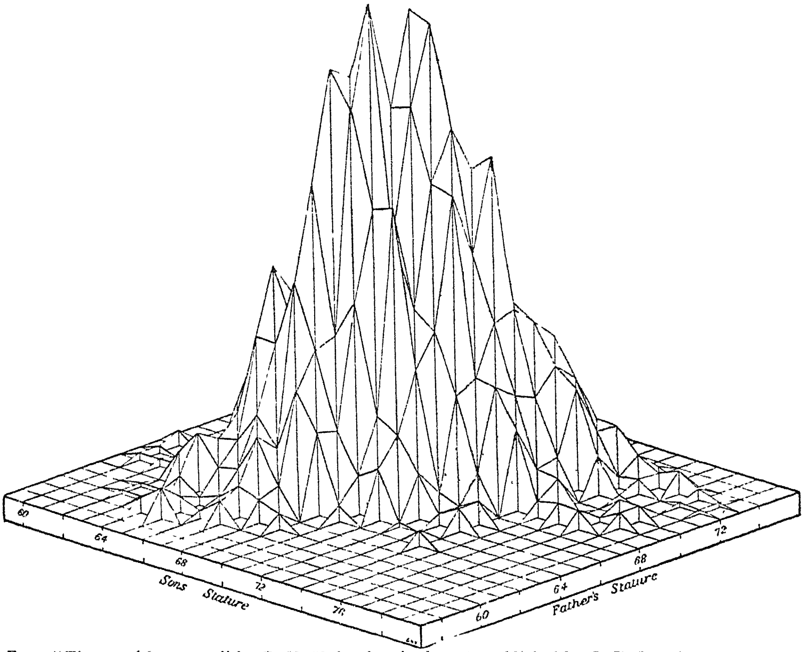
To the x-axis, or distance down the length of the room, let us give the values of time, letting the first unit of distance represent one year, the second the next, the third the following year and so on, so that along the length of the room we have, on the x-axis, a scale of years. And at each year we notice the cross-wise lines and the vertical lines are repeated to form a co-ordinately ruled plane perpendicular to the x-axis. To the cross-wise distances of the floor along the y-axis let us give, for example, a scale calibrated to tens of dollars, and running, let us say, from zero to two hundred dollars, there being twenty cross-wise divisions of measurement. To the vertical lines or the ordinates parallel to the z-axis, let us give a scale calibrated in hundreds from zero to one thousand, there being ten vertical units of measurement. In short, to each of the three dimensions or axes of the cubic volume of the room, we can attach scales of calibrations similar to the scales in ordinary curve-charts along its two dimensions or axes.

Returning to the end of the room to the vertical plane cutting across the room at right angles to the x-axis and intersecting the x-axis itself at the point of the first year, let us chart upon this plane a frequency curve showing the number of sales of various sizes for the concern whose business we are analyzing for the year indicated on the x-axis. Then on the next plane, intersecting the x-axis, the point on its scale for the next year, let us plot a similar frequency curve for the next year. And so through all the planes, let us plot frequency curves, one for each year upon the plane intersecting the point of that year upon the x-axis. Let us plot these curves by attaching red strings to the network of strings which makes up the planes.

Having completed the series of curves for all the years, we can step off and look at the result. The curves are likely to show great similarity with but slight changes in their exact positions from year to year. These changes of the positions of the red-string curves show us the changing nature of the sizes of sales made by the house. The series of red strings seems to outline a billowy blanket or irregular curved plane. If through the series of curves we should attach similar red strings running lengthwise down the room, connecting corresponding points upon the curves, this blanket or suspended irregular surface would be more visible. Examining any one of these new connecting strings, we would find that each one of them forms the curve of the historical changes in the number of sales of each size through the various years. In short, the blanket could have been made by plotting the various historical curves along the room rather than the frequency curves across the room. The blanket itself is in fact nothing but a historical projection, carrying a frequency curve through a number of years, and yet exhibiting all its points at any point of time.

It should be remarked that the name "double frequency series" given to this type of data is a loose one, used to describe both truly double-frequency series and historical-frequency series. The curved plane which we have just plotted obviously represents a historical-frequency series, that is, the data of a frequency series carried through a number of periods or points of time. The double-frequency series proper is similar in all respects, except that time is not one of the independent variables. In the double frequency series proper, the data of a frequency series is carried through a number of changing

conditions or classifications which, like time, form a connected or variable series. The distinction between the historical frequency series and the double frequency series is not important, either in computing or charting.



From "Theory of Statistics," by G. U. Yule (fourth edition), published by J. B. Lippincott.

Fig. 488. Smoothed Frequency Surface.

Showing the correlation between the height of fathers and sons.

As you will see, the name "curve" is really a misnomer for this form of chart, for the chart does not merely exhibit a curved line, but exhibits a series of curved lines forming a curved plane. The word "surface" is ordinarily used for this form of chart, though it might also in a loose sense be called a curve. It is not always true that the succession of curved lines or true curves can be joined to form smoothed surfaces. It may be that the phenomenon requires that the joining be made in staircase fashion. The same consideration applies to the smoothing of the surface between plotted points as applied in the smoothing of the ordinary curve. It may even be desirable to smooth the curves along one axis and leave them in staircase form across on the other axis. The familiar and most useful forms of the curved surface are, however, the smoothed sur-

face (smoothed along both independent variable axes), and the double-frequency polygon which is in staircase form along both axes.

WET AND DRY MONTHS Summary of the number of times each month has been first, second, third, etc., in order of humidity during the years 1868 to 1906, 38 years in all. Taken from flow of Croton River, N. Y., at dam.

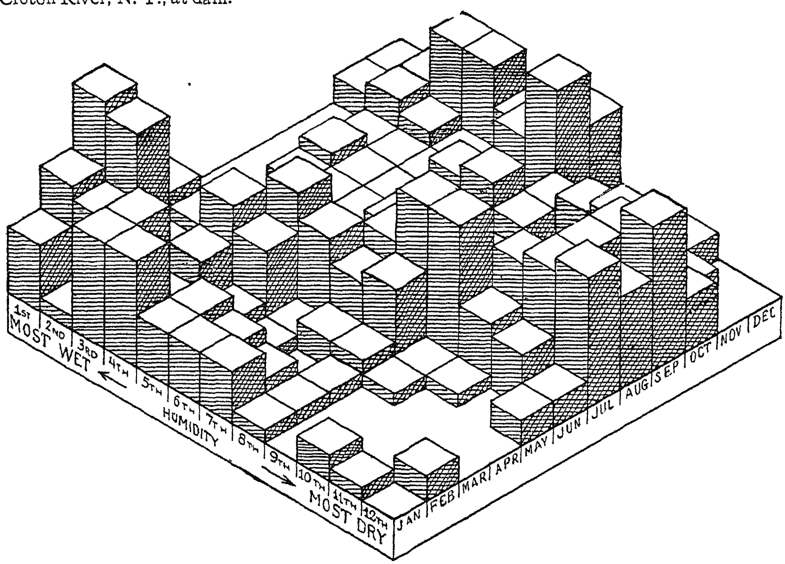


Fig. 489. Staircased Frequency Surface.

To make a physical model of these two-dimensional curves is not difficult, but is rather tedious. Different methods have been recommended when the solid model is desired. A simple practice is to cut strips of wire and mount them vertically upon a board, the lengths of the wire emerging out of the board representing the z-ordinates, each wire being cut off at the point where it would intersect the curved plane or the individual curve. If the board has been previously drilled with holes to admit the wires, the holes being at the intersections of the two sets of abscissae, it is not difficult to erect in a very short time a forest of these wires, their top ends readily outlining the contour and shape of the curved plane. The next step is to place the board with its wires in an enclosed box and pour enough plaster of Paris over it to cover all the wires. The last step is to cut away the plaster of Paris until the ends of the wires appear, the plaster of Paris being easily scraped off so as to form a smooth plane. It is also a convenience to out-

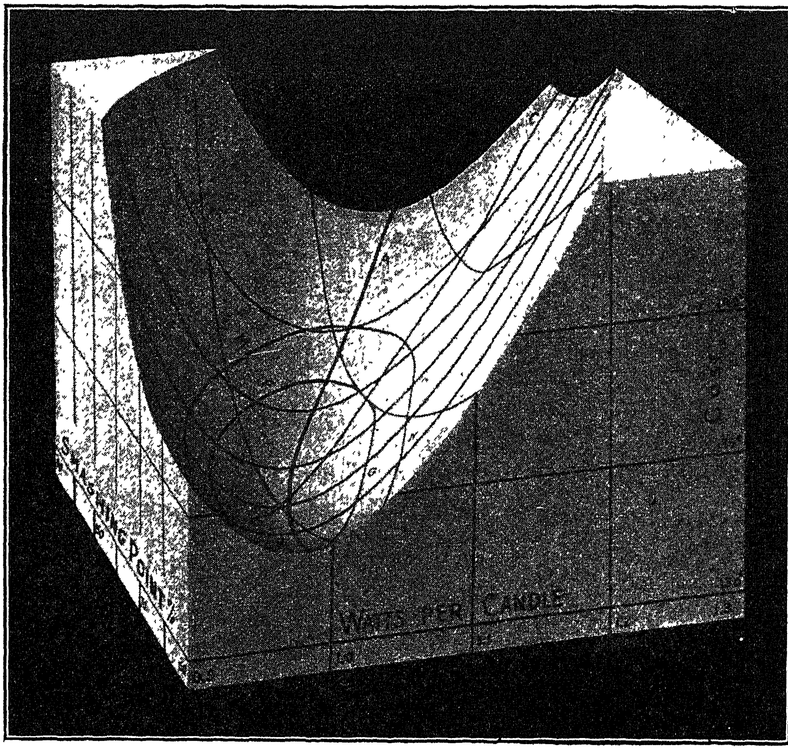
line the various horizontal co-ordinates upon the sides of this solid model and upon the curved plane at the top of the model, where these horizontal co-ordinates intersect the curved plane. These co-ordinates can be shown by thin black lines thus facilitating interpolation and the reading of plotted values from the model.

Another method by which the same kind of solid model can be obtained is to plot the individual curve upon pieces of stiff paper, one complete set of the curves being plotted so that they can be set up side by side in the same way that the strings were attached to form vertical planes in the network in our imaginary room. When the variations of these different curves are great, it is also of advantage to plot another complete set of curves of the same data upon the other independent variable, this second set of curves showing the appearance of the connecting strings finally added in the imaginary room above discussed. In this case, we have two complete sets of curves, one for the longitudinal curves and the other for the latitudinal curves. By cutting slits in the heavy paper as described in the last chapter, the two sets can be fitted together as the divisions in the ordinary egg box in which you buy a dozen eggs.

If the paper upon which the curves are plotted is cut away at the curve, that is, if you take a pair of scissors and cut through each plotted curve, the lower halves of the chart will have the outlines of the curve for their top edges and when they are joined together like an egg box, they will form a three-dimension model which can be collapsed, if a collapsible model is desired. If a rigid solid model is desired, you can pour plaster of Paris into this collapsible model, scraping the surface of the plaster of Paris down to the edges of the paper curves so as to obtain the smoothed curve plane.

When the curved plane is to be in staircase form, it is easily built up out of blocks of wood. The procedure here is very simple. You need merely take a long piece of finished lumber with a square cross section and cut it into strips the same length as the wires which you would have left standing in making a plaster of Paris model. The wooden rectangular cubes are then glued together to form a single solid model. There is no need of projecting the ordinates upon this model, as the edges of the individual pieces of wood indicate the ordinates, but it is well to mark in the horizontal rulings around the side of the

model so that the height of the individual pieces of wood can be easily estimated from an examination of the model. Needless to say, in all types of solid models the scale calibration for



From R. E. Scott, in Harvard Engineering Journal, by permission.

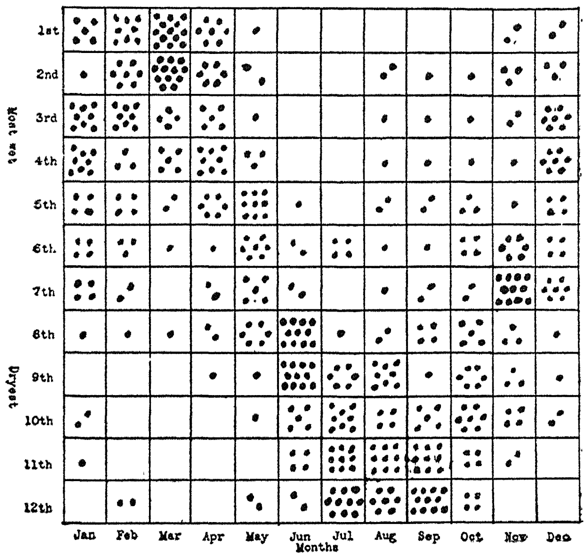
Fig. 490. A Solid Model—Rounded.

Model showing cost of light in cents per 1000 Candlehours with 40-watt "Mazda" lamps, for any combination of efficiency and smashing point, where price of lamp is 50 cents and of current 10 cents per Kilowatt hour.

the horizontal distance should be around the edges of the bases of the models.

Both the smooth and the staircase curved planes can be often easily pictured upon isometric or other axonometric paper as described in the previous chapter, eliminating the cumbersome and tediously constructed solid model. In general the staircase form of curve-plane is perhaps more easily projected upon this paper than the smoothed surface. The paper has the disadvantage of presenting the view of the

model from one side only. Therefore care should be used in the selection of the side from which the model will be seen on isometric paper, in order to get as much detail as possible, that is as many parts of the curved plane visible as possible upon the isometric drawing. The isometric drawing is perhaps best adapted to symmetrical forms of double frequency series or to the double ogive or cumulated double-frequency series, for in the case of ogives the variation can all be seen from one side anyway. A third dimension can of course be symbolized upon the ordinary two-dimensional chart by the use of colors or shading which epitomize the staircase form of



WET AND DRY MONTHS
Summary of the number of times each month has been first, second, third, etc., in order of humidity during the years 1868 to 1905 - 38 years in all. Taken from flow of Croton River, N.Y., at dam.

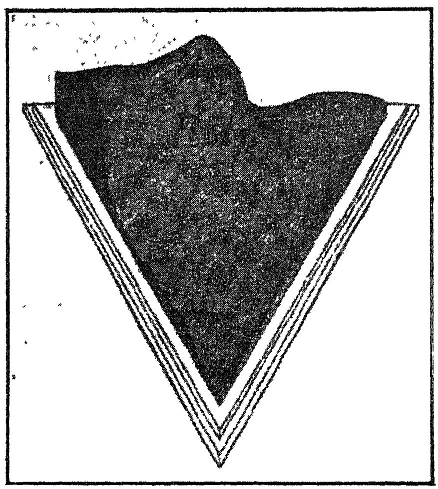
Fig. 491. An Orthographic Model.

curved plane, or by orthographic lines (similar to contour lines in topography, or to isothermal lines in weather maps) which zone off the smoothed curve plane.[1]

In general the frequency surface has for its two horizontal dimensions, that is, the axes of its two independent variables, a rectilinear pattern of co-ordinates. It may be, however, that upon these co-ordinates a series of irregular shaped out-

[1] See Chapter LIV.

lines are traced, which are the boundaries of irregularly shaped areas to which our dependent data (the function or body of the tabulated double-frequency distribution) attaches. The map is a special case of this in which the horizontal co-ordinates mark off longitudes and latitudes and the areas represent geographical localities. The orthographic chart is a general case of this in which the irregular outlines are called contour lines and the areas are merely zones of equal functional value. On the other hand, it is not necessary that the horizontal co-ordinates be rectilinear to begin with; thus we may use trilinear co-ordinates, such as are used in the hundred per cent triangle, for our horizontal plane or base. Indeed when the 100% triangle is used we have a peculiar chart showing four



From John B. Peddle's "Construction of Graphical Charts," by permission.

Fig. 492. A 100% Triangle Model—Four Variables.

Professor Thurston's solid tri-axial model showing the efficiency (by height) of alloys of three metals in various proportions.

variables, three of which combine by addition (or logarithmically, by multiplication) to form a constant.[2] The solid built up from a 100% triangle, might by its altitudes (or z-ordinates) show, for example, the efficiency of foods whose composition is indicated, as to fats, proteids, and carbohydrates, for example, by position horizontally. Various pro-

[2] *Cf.* Robert Thurston, *Glyptic Models*, in the *Transactions of the American Society of Mechanical Engineers*, 1898.

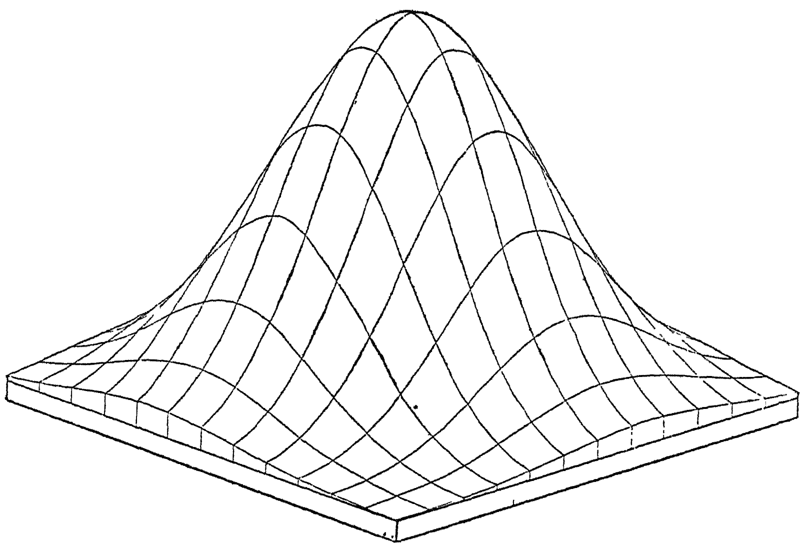
jections of the horizontal scales (for the independent variables) are possible, including even a probabilities or double-probabilities projection. These obviously have but limited usefulness, in more complex cases of equating the phenomena with these variables, the probabilities projections being designed to convert the ogive surfaces into flat tilted planes, and the other projections having the same objects for other series, all with a view to the writing of equations.

It is difficult to adhere strictly to the field of chart making in these more interesting and important mathematical charts. We are constantly tempted to wander off into the field of statistical methods with which these charts are sometimes intimately connected. And though this book is not a manual of statistical methods we shall here digress long enough to sketch in a few of the more important uses of the double frequency curve. For in the statistical laboratory the frequency surface is often used for the study or presentation of correlation and association between different methods of classification of the same phenomena. Correlation between two historical series can, as you have seen, be best shown by the juxtaposition of their rate-of-change (i.e. logarithmic) curves. But correlation between two independent variables of the same data can be shown in detail by the stereographic or three-dimensional model.

The nature of the normal curve of error, that is, the dispersion about a central or most typical point, which is to be expected under the operation of the law of chance variations, has already been explained in so far as variation along one dimension is concerned. But variation can equally well take place along both dimensions when the nature of the phenomenon is such as to permit it. Thus the dispersion of gunfire from cannon varies as to both distance and direction (range and deflection). The coaction of the same probable dispersion along both axes, that is both longitudinally and latitudinally, results in a cone-like peak whose sides appear to slope along the curve of normal error, when seen from any side.

When the double-frequency curved plane presents this form, the thought is naturally suggested that the two independent variables upon which the grouping or classification of the functions depends, are really independent in their action, and do not affect each other. When, instead of a cone with its sides following the normal curve of error, we have a ridge diagonally

across the chart, whose cross section may or may not resemble the curve of error, we have a rough means of measuring the correlation between the two bases, the very narrow ridge presenting high correlation and the wide-spread irregular ridge showing low correlation.



From G. U. Yule, "Theory of Statistics," fourth edition, published by J. B. Lippincott.

Fig. 493. The Normal Frequency Surface—Rounded.

The two-dimensional curve or curved plane, either smoothed or stepping, is perhaps less often used in business than it ought to be. It gives to important data a valuable projection, changing through time and conditions in the case of historical frequency series, and illustrates the co-action of the two independent classifications or changing conditions in the case of the double-frequency series. The labor of preparing the charts is not great and the illuminating pictures they present are ample recompense.

Chapter LVI

RELIEF MAPS

It is in the more elaborate form of maps that we find the most frequent and perhaps the most generally understood form of three-dimensional charts. Every one is familiar with the relief-map model used in the school room, in which the two horizontal dimensions are used for the latitude and longitude as in the ordinary map, but in which the varying heights or altitudes of the model indicate the altitude of the land, mountains being shown by ridges, and rivers and valleys by cuts and hollows.

Business men and economists have perhaps little interest in the physical contour of a country, but the principles of the relief-map can be used to illustrate a large variety of other things than the actual height of the land levels. The sales manager may be interested in a relief-map in which the z-ordinates, that is, the vertical distances or heights of the relief-map model indicate the density of sales, as shown by per capita sales, per dealer sales, or by other means. The engineer may be interested in a relief-map in which the height indicates the amount of natural resources, water power, mineral deposits, and so on in the various localities. The economist may be interested in a relief-map showing the financial resources, wealth, crop yields, or other sociological conditions in the locality.

The usual way of presenting these relief-maps upon flat surfaces is to indicate the height which the actual relief model would have in its various portions, by different kinds of colors or shading. If colors are used, they should be arranged along a color chromatic scale so that the colors themselves, by their changes (for example, from red to blue, through orange, yellow, yellow-green and blue-green), have a natural significance and can easily be understood. If the colors are carefully chosen for their tints and intensities, they can be

successfully photographed, and will show on the photograph as black for the reds and white for the blues, and varying through dark grays to light grays for the intermediate colors.

The problem here is to secure color tints which have not increasing optical intensity but increasing actinic intensity, for the camera does not photograph different colors precisely as the eye registers them. Many attempts have been made to adopt scales of increasing intensity of color, regardless of their chromatic sequence, but these attempts are almost always unsuccessful because of the extreme variation of available colors which may be used with apparently the same optical results, but with actually different actinic or photographic values. Moreover, the arrangement of colors solely according to ocular density or intensity is unsatisfactory because for the significance of each color the reader of the chart must refer to a key.

There are many different ways of applying color to charts. The disadvantage of water colors is that they tend to run upon the paper and are difficult to shade off, no two mixtures being precisely alike when laid at different times on different charts. Colored water-proof inks, ordinarily used in drafting, must often be diluted or their intensity will be so great as to hide any printing or labelling which was intended to show through the color on the chart. For extreme transparency, photographers' Japanese transparent inks and lantern-slide colors can be used. Perhaps the best results are obtained in general with ordinary wax crayons, the cheaper and more waxy they are, the better. These can be laid on very thickly and evenly, and can then be scraped away with a sharp knife edge, leaving a delicate tint through which all printing and labelling will be easily seen, and not warping or wrinkling the paper.

Shadings are of many different kinds. The Census Bureau makes great use of dot shading, a number of small dots being placed upon the paper, scattered over the locality and by their number, showing the values of the data for the locality. The method has a decided advantage in that the charts tend by their crowding or scattering to indicate density visually, but it is a tedious method to follow in the making of the chart and the results do not afford any degree of accurate reading. It is impossible for the average reader to count the number of dots where these have been thickly placed and where they are so thick as to form almost black areas the significance of the

dot has become entirely pictorial. The method is, however, far better than another one which resorts to dots or circles of various sizes in which the areas inscribed in the circles indicate the values of the data. For in these circles of various sizes, we meet with optical illusions and difficulties of accurate chart-reading as well as chart making, described very early in this book. The area of the circle although a two-dimensional measure, is used to illustrate one-dimensional data. Dot-maps are often most easily made with colored map pins or map tacks which have been described in an earlier chapter.

Much the better form of shading is secured by a careful scheme of hatching and cross-hatching lines so arranged as to give an optical effect of shading from white to black and at the same time sufficiently different in pattern to be easily identified from a key or appended scale of shades.

Several such patterns of useful shadings have been designed. The work of drawing in these shadings, however, is sometimes very great, it is difficult to rule these hatching lines uniformly without a special instrument (section-ruler) and the whole process takes a great deal of time. It is therefore sometimes better for the average chart which must be prepared in black and white, to mix India ink and Chinese white in various degrees and ink them onto the chart as so many tints of gray. When this is reproduced photographically the effects are entirely satisfactory for photostats, but are useless for blue-prints. As these tints require half-tone engravings for printed reproduction, it is more convenient, when the map is intended solely for printed reproduction, to use the forms of cross-hatching and shadings known as "Ben Day" in the printing office. When Ben Day is used, your original drawing need have no shading at all, the various types of Ben Day merely being indicated by numbers or symbols in blue pencil on your drawing; the engraver will insert them properly.

Relief-map models may be compared to the smoothed-plane curve of the last chapter, in that the changes of types are never abrupt but are gradual. The staircase form of these map models is an elevation map or table-land map which is less often seen. It can be best constructed with a few sets of maps, mounted upon boards of different thicknesses, and cut or sawed apart along the State boundaries. Children's puzzle toys are sometimes made in the form of maps of the United States in which the individual States have been cut away in

this fashion, and may prove useful for this purpose. A picture of the elevation or table-land map is, however, easily drawn upon isometric paper, or upon plain paper, by tracing the State outlines from a regular map of the United States, after shifting the position of the tracing paper slightly to correspond to the representation of height or elevation of each State. Such maps are very effective in their way.

One disadvantage of the colored or cross-hatch map upon paper drawings is that it represents a staircase form of map while most phenomena should really be smooth, as the transitions from State to State are not abrupt but gradual. To meet this problem, the colored or shaded map can often be skilfully converted into a zone map, in which the colors or shadings have been zoned so that no two colors or shadings appear side by side upon the map except those which are consecutive in the key or scale of colors or shading. Where the data applies to the entire State or other territory, these zonings are of course arbitrary and tend to alter significant areas in precisely the same way as the smoothing of a frequency polygon destroys significant partial areas of the polygon. The zoning should therefore be made very narrow, along the outlines of the States, in order to leave as large as possible a portion of the space properly colored according to the data.

However, if the data does not represent the entire State or other territory, but merely indicates conditions at certain points, such as certain cities, the zoning should always be done and the zones should be of equal width between any two observed points. A familiar example of this type of zone map is to be found in the map used by the Weather Bureau showing high and low pressure areas from day to day and temperature lines across the country. In fact the entire zoning process is merely an attempt to reproduce for the particular data, the same excellent results achieved by these isothermal lines upon the weather map. The contour lines upon topographical maps are examples of the same type of zoning or orthographic rulings.

The technique of these various presentations of the third dimension upon the map is fully described in a previous chapter and we have here only hastily recapitulated the more common forms of maps. In the case of all the maps so far considered, the reader will notice a limitation, in that each map is capable of presenting but one set of data. Two figures for each locality cannot be shown upon the same map. There are, however,

two ways in which more than one geographical distribution can be shown upon the same map. The first of these methods is the very obvious one of combining on the same map two methods of showing the third dimension. Thus on a flat map both colors and cross-hatchings can be used simultaneously, the colors to show one set of variables; the hatchings, another. The results are not wholly satisfactory. If bars or other area charts (circles, stars, etc.) are used to show values of course a series of bars or areas can be used in each part of the map, corresponding bars being perhaps distinguished by color or shading. If a stereographic map (i.e. solid model or axonometric drawing) be used, of course one set of values is shown by the altitudes (z-axis ordinates) and another by colors or shadings drawn upon the resulting surface. And always the use of numbers actually entered upon the map may give us further values not graphically displayed. All of these methods give us what might be called multiple maps, in that they are combinations of two or more map surfaces. The second way of showing more than one geographical distribution is more laborious in construction, but also much more illuminating. It is the method of the bead-map. It gives us not only multiple maps, that is combinations of distinct distributions, but also it gives us compound maps, that is segmentations of a single distribution. This last feature is one which cannot be satisfactorily achieved by any other graphic methods and can be shown only on a flat map by figures. The bead-map must be classed with solid models, for it is a rigid body occupying space in three physical dimensions as fully as if it were of plaster-of-paris, wood, or some other substance. In its construction we fall back upon the upright wires which were used in the construction of the plaster-of-paris model.

The bead-map takes its name from the fact that after properly cutting the vertical wires, it is customary and most satisfactory to string beads upon them before their ends are inserted in the map. Beads can be obtained for use in this way, in many different colors, at the average department store. Sometimes glass beads of uniform sizes, especially advantageous for this work, can be obtained from the publishers of charting material. The wires themselves should be fine, of medium strength and spring and can also be obtained from charting-material publishers, especially adapted for this work. Where only a few beads are to be strung and the wires do not

extend very far above the paper, very long and thin steel pins such as are used in natural history museums, can be used for the wires, the heads of the pins holding the beads and preventing their escape. When wires are used, small knots must be tied at their upper ends, to prevent the beads from coming off the wires. The beads on each wire should be all of one color except that every tenth bead, or (according to the scale for beads), every significant bead, should be of a contrasting color so as to facilitate the counting of them by the reader of the map.

These maps can be made with as many as three or four sets of separate wires for each State in the Union, each set representing a certain figure or set of data, on maps which themselves are no larger than ordinary-sized letter paper, that is, $8\frac{1}{2}$ by 11 inches. The different wires for each State should be placed upon a single line or row across the State so that they can be easily compared and form, as it were, a vertical bar-chart upon the State.

The variety of purposes to which the bead map can be put is as great as the uses of other relief or color maps. The sales manager, for example, will be interested in a map in which there are four columns of beads in each State, a column of green beads indicating the population or potential market in each State, a column of red beads indicating the sales of his competitors in the State, a column of blue beads showing his own sales in the State in the previous year, and a column of black beads showing his sales this year. A comparison of the red and black beads shows him how well he is keeping up with his competitors, while a comparison of the green and black beads shows him how well the market is being saturated by his goods, and a comparison of the blue and black beads shows him his last annual increase or decrease of sales.

Another convenient form of this map is sometimes called the "tree-map." In this the stringing of beads upon wires has been eliminated entirely and small pieces of colored wood sticks substituted for them. A full equipment of wood sticks of uniform thickness and shape, but of different lengths and colors, can be obtained from the manufacturers of kindergarten toys, being often sold for kindergarten work. When these sticks are being used, their ends, to be inserted in the map, should be sharpened with a knife so that they can be easily inserted, after first marking upon each clearly the dis-

tances which should be left exposed, sticking up out of the map. After they have been driven into the map to the proper distance, they should be removed and into the holes made by them drops of glue should be placed, the sticks being then replaced in the holes and allowed to set in the glue. The tree-map does not afford the possibility of exact reading which was possible in the bead map, where the beads themselves on any wire could be readily counted. If this feature is desired, small colored bands should be drawn at the points about the sticks of wood at the heights where the distinctly colored beads would appear, marking off on each stick of wood the ordinates of the various convenient values on the z-axis.

Needless to say, the map for beads and wires, or for sticks of wood, should be mounted in the same way as pin maps, which have been described in an earlier chapter. They should have at least three layers of corrugated pasteboard under them to hold and protect the ends of wires or sticks.[1] Neither bead-maps nor tree-maps are convenient to file or to have about in large numbers as they are apt to get damaged. The best way to file them or to carry them about is to use small wooden boxes or cases into which they can be slipped easily and fit compactly, and in which they are prevented from moving about by small retaining flanges inside the boxes or cases.

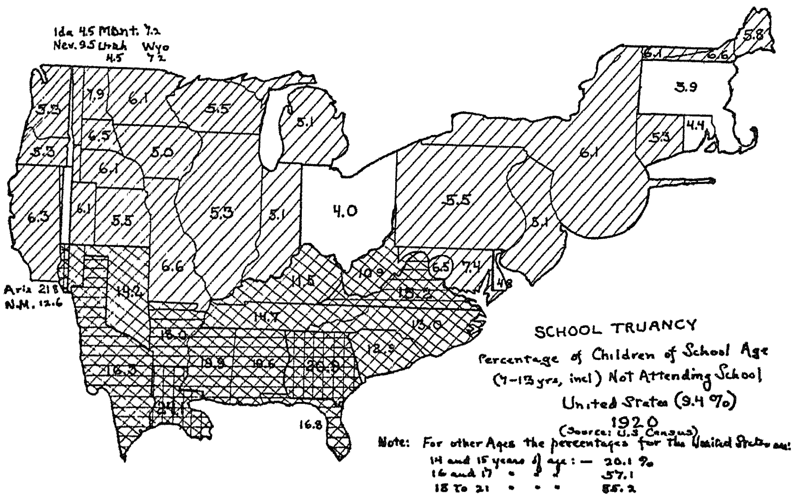


Fig. 494. Cross-hatched Map on the Population Projection.

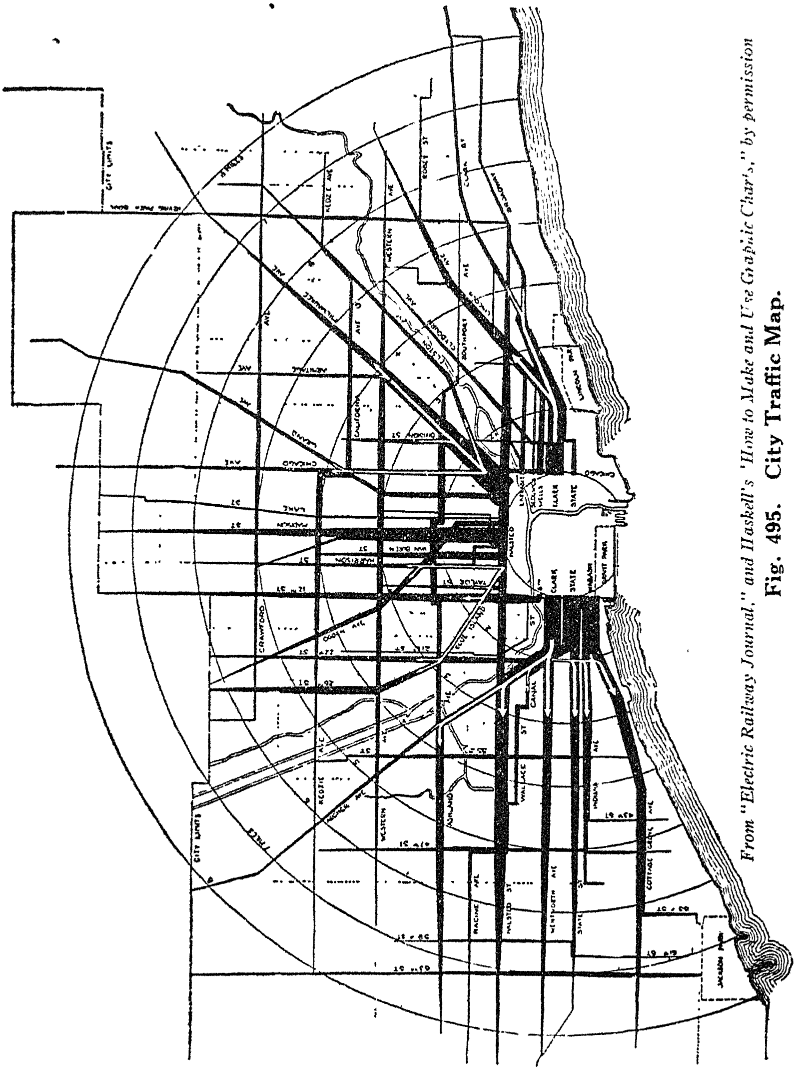
[1] See Chapter IV.

Keys should always be provided with every map, to explain the significance of colors, shadings or beads, and wooden sticks. These keys serve the same purpose as scales in curve- or bar-charts. They should be complete and carefully worded. It goes without saying that the special projection maps described in the chapter on population maps, can be used in the place of the ordinary land-area map for all cases where the significance of map areas is better shown by such projections.

Maps need not be used only for the display of characteristics of entire localities and territories, but can also be used for the analysis of routings and conditions along certain routes or at certain points on the map. In this case we are not concerned with areas on the map but with lines or points upon them. The use of strings upon maps, connecting map pins or map tacks, has already been described in an early chapter. These strings can be used not only to indicate actual routes which will be followed by sales managers, travellers, or traffic, but can also be used to indicate spheres of influence, authority, or other connecting influences. Thus the circulation of a number of newspapers situated at different points of the country can be shown by strings (or indeed by mere ink lines) radiating from their places of publication to the residences of their furthest subscribers, different colored strings (or ink) being used for each newspaper. Such a map is sometimes useful in the analysis of newspaper circulation for advertising campaigns. Likewise the line of authority from central office to the various branch houses and from branch offices to the individual agencies can be similarly shown by radiating strings. Such maps properly belong to the class of combinations or superimpositions of route-charts upon maps described in the chapter on combinations of non-mathematical charts.

When, however, we attempt to show the volume of traffic or travel, or the extent of any other connecting phenomena between two points upon the map, we come quickly into the field of three-dimension maps. If we wish to present this graphically, a very effective method has been found of preparing ribbons of stiff, colored cardboard or paper, and mounting these ribbons on edge along the route or line of traffic or connection, so that the height to which the ribbons rise, will indicate the volume of traffic or other connecting phenomena. The same result can be more easily obtained by the use of

colored strings connecting columns of beads which have been previously erected at the points to be connected. The strings



From "Electric Railway Journal," and Haskell's "How to Make and Use Graphic Charts," by permission

Fig. 495. City Traffic Map.

are easily tied to the columns of beads and run back and forth, one string between each layer of beads, forming fences similar to the old-fashioned rail fences and indicating by the number of strings or rails the volume of traffic or other figures for the connecting phenomena. If we wish to present the three-

dimensional data upon a flat surface, for convenience in handling and filing, and the number of routes or connection lines is not great, we can show the comparative height of the various ribbons or set of strings by colored or shaded bands drawn upon the map connecting the points which the bead or string fences would connect. In this case, the widths of the bands representing the volume of traffic or other connecting phenomena.

The reader will have seen by this time that map-charts are almost a field of charting in themselves, with wide diversity and flexibility and an infinite variety of forms, capable of the widest variation and adaptation for special purposes. In fact, the map can be considered as distinct from all other types of charts in that the fundamental two dimensions, that is, the two dimensions of the base-map itself, are used solely for the purpose of displaying geographic position and location (except in population maps) and not for strictly mathematical relations and that the mathematical relations must be charted upon this ground-map by the use either of a third dimension or of superimposed drawings representing a third dimension. In a sense therefore, the map can be considered an inefficient or wasteful type of chart. For economy of space and charting dimensions, the superior form of chart for all data having a geographical basis is the bar-chart in which the geographical location is shown by a list of stubs and the independent variable, that is, the geographical location, occupies only one dimension on the paper. This applies not only to maps, but also to diagrams, floor plans and other illustrations of space or physical localities, all of which can be treated in the same way as maps have been treated in this chapter. But in spite of these disadvantages and inconveniences, the map is so useful that it can be strongly recommended to the studious chart-maker as a powerful method of displaying such facts as are of sufficient importance to justify the greater labor and care involved.

PART VII. CONCLUSION

CHAPTER LVII

THE STATISTICAL MATERIALS

Few, perhaps, of our readers, have run the gamut of chart forms and methods which we have described in this book without realizing that there is almost as surely a natural evolution in charts as there is in other sciences or arts. It is possible and would indeed be interesting to construct a diagram in the form of a tree-chart, showing the development of each chart form out of common root-forms. Within the bounds of our limited ability we have constructed this book after such a pattern. And as time passes and new forms are invented, or new modifications are introduced, it will always, doubtless, be possible to relate them to existing forms and allocate each to its proper niche upon such a diagram.

But more interesting than the classification of chart-forms is the classification of the statistical materials which they illustrate. For the numerical arrays and tabulations, the counts, samplings, enumerations, and reports which we call statistics present even greater variety and heterogeneity than the charts by which we may picture them. Nor need such a classification be wholly academic. The coding and systematizing of graphic methods can hardly progress far without becoming tangled in the chaos of statistical forms, a confusion from which it cannot again escape until we have set to order the statistical stock-room.

If then, we could succeed in so neatly classifying and pigeon-holing each type, species and hybrid form of statistics, that the novice could readily identify each specimen, we would set for ourselves this aim: That each variety of "statistic" should be clearly labelled and marked with the one, two, or more ways of charting suitable for its illustration. We would have this code, key, or system, so simply set forth that the economist and the business man, be he ever so untutored in the science, could easily locate in it his particular bit of statis-

tics and as quickly set out effectively to chart it. This we say would be our ambition, were it possible. And while many may doubt its possibility, yet in this chapter we shall venture a few first steps in its direction, only bespeaking in our readers a tempering of judgment with generosity for the short-comings and failures which attend us.

We do not progress far in statistics without noticing that all numbers are purely adjectival, and that to each number, in order that it may have a meaning, a substantive must be attached. If we only make mention of so small a thing as two pins we may observe that the numerical adjective "two" holds meaning only when attached to the noun "pins." Speak of three needles and we note that in this beginning of a statistical collection we have changed both adjective and noun. To make it look quite professional and uninteresting we should tabulate it, thus:

Pins	2
Needles	3

But suppose that these pins were ordinary household pins, while we may discover elsewhere five safety-pins, to be added to our collection of pins, bringing it up to seven, thus:

Pins, common	2
Pins, safety	5
Needles	3

Now we notice that while the noun has remained the same, another adjective has been added beside the numbers. Many such qualifying additions could be made. And it is the object of this homely illustration to point out that while numbers are always adjectival, not all adjectives are numerical, and that numbers like other adjectives, such as "safety" or "common," have a meaning only when associated with some substantive.

But we stand too long on pins and needles. Let us only keep in mind the point that in all statistical work we must early recognize two variables, variates, or variable facts, which for convenience we have always distinguished as independent and dependent. It may be that at times one will seem the noun and the other the adjective, or that various adjectives will vie with each other for importance, and that at other times in the same data a reverse arrangement will be more useful to us, but always at one particular time we must use one variable as independent, that the other, clinging to it,

may be dependent. And statistical technique is largely a matter of the proper marshalling, commandeering and buffeting about of these two sets of variables until they behave in a way that yields up to us an intelligible message. There is no hard and fast distinction between the two variables and their degrees of independence or dependence—they are truly interdependent—but always there are hard and fast rules which govern the treatment of them in such a way, that whichever variable is playing the independent role may be given such and such handling, and the other which is playing the dependent role has such and such other possible operations. As a rule we usually let the independent variable take its own course and put most of our efforts upon the dependent one, but this is not always so.

Another thing which we may note at once is that statistics may come to us either in singular or plural form. We may have, as it were, a single "statistic," or a collection of statistics. Thus the simple fact:

Pins..............7

may be the alpha and omega of our desired information; or the more elaborate statement:

Pins, common.....2
Pins, safety.......5

may be our objective. Here we come upon the distinction between what may in a wider sense be termed "averages" and "distributions." Commonly, this particular average would be called a "total," for it is the total of its parts, which parts are forcibly brought to mind by the subsequent distribution. And in a specialized sense, averages and totals are very distinct. Thus we would say:

Pins, total.........7
Pins, common.....2
Pins, safety.......5
Pins, average......3½

But in a wider, perhaps more precise sense, all so-called totals are merely averages—if you will, averages of the counts taken of the items, or averages for particular selections or samples. Into the various kinds of averages we shall not delve—their consideration forms the large part of most elementary treatises on the science of statistical methods. Let us compromise by distinguishing at once single and collective statistics as "averages or totals" and "distributions."

Now before going on to collective statistics, or distributions, let us look just a little longer at the single entry or item:

Pins..............7

We note that the number, 7, is a count. But a count of what? A count of the number, you say, of whole pins. So it is. And being so, it is so obvious that it has not been included in the statement. Could we count these pins in any other way? Suppose we count the half-pins, then there are fourteen in all. But this is quibbling, you say. Very well, suppose I inform you that these pins are exceptionally rare and are valued at a dollar apiece. Now you can count them as:

Pins...........$7.00

or more fully:

Pins (value) 7 (dollars)

Again we discover that they are railroad coupling pins and weigh two pounds each. Now we can write:

Pins (weight) 14 (pounds)

And all the time we are merely describing in different ways the same objects! In short, in statistical work, we deal with various numbers and by them we count various items in terms of or by means of various units of measurement. The units may be, roughly, measures of volume or of value. Volume statistics refer to physical volume in a loose sense and may be in terms of linear, surface, content, weight, or other measures, including the mere count of the number of items. Value statistics refer to intrinsic worth as indicated usually by actual or potential price or cost, and are generally in terms of money of one currency or another.

While the separate or isolated average or total may appear in any of these forms, it does not in any of them become a proper subject for a chart, for a chart is only useful as a means of comparing two or more figures. We do not enter the field of charting therefore, until we come to collective statistics or distributions. In turning to these, however, we must keep in mind the various possible forms of the individual item or "statistic" for it is obvious that they apply as well to the collection. The collection or table of items, then, will contain both independent and dependent variables, and will be composed of either averages or totals (now become sub-totals), measured in units of volume or value. In fact, the collective statistical statement often appears to be no more than a mere agglomeration of individual statements placed together upon a page.

At other times the collective statement appears to be a more detailed elaboration of some individual statement. In the latter case the name distribution is clearly called for, but it is also true that the most patently heterogeneous aggregation or conglomeration of figures can generally be regarded as a detailed distribution of something or other. In this sense we may speak of the simple individual item as "undistributed" and the collection of items as "distributed."

It is in studying the various types of distributions that we strike the first important distinctions of data, and, by corollary, of charts. Four types at once come to mind, which, for convenience, we may call respectively, abstract, geographical, frequency and historical distributions. Indeed it is not improbable that in the course of time the compilers of statistical volumes will adopt distinct tabular forms for each of these types and consistently maintain these forms in their compilations. The Bureau of the Census has already adopted a more or less standard form of table for geographical distributions where these cover the States of the United States. Even more successful is the excellent standardization of historical tables by the same bureau for its "Survey of Current Business." By the side of the usual confusion of tables in most statistical compilations the simplicity and clarity of these forms is indeed refreshing.

There is very little overlapping of the four types of distributions. The first two are essentially logical, the last two are essentially numerical, in respect to the bases of their distribution or classification, which form the independent variables or "stubs," in them. The first and third are generalized types, the second and fourth are merely extremely common and important species of the first and third. The basis of the second, the geographic distribution, is space, while the first, the abstract distribution, can have as a basis, any other set of logical relations. The basis of the fourth or historical series is time, while the third or frequency series can have as a basis any other set of numerical relations. And, of course, we can have what might be called composite statistical tables presenting two or more of these distributions simultaneously. Indeed, in most statistical compilations the composite distribution is chiefly used for its convenience of comparisons and economy of space. Not only can two abstract distributions be made to interlock in a single composite one, but two geographic and

two frequency distributions are often so combined. And there can be any combination of different types. In the present book the distribution placed at the sides of the table has been called the series of stubs, generally considered the more important, and the other, placed across the top of the table, has been called the series of column-headings or captions.

If we should attempt the explicit description of statistical distributions of these various types—confining ourselves to single distributions only, since what applies to these applies likewise to each of them in composite distributions—we would find certain salient points which must be noted about the independent variables in each type. Thus in describing (or cataloging) an abstract distribution the important things to note are the basis of the distribution (whether it is nature of diseases, causes of accidents, kinds of articles, races of the population, sex, marital condition, or what not) and the number of items in the series. Little more can be done to categorize the abstract distribution. And we may note that in graphics, the abstract distribution is amenable to no connected method of illustration, such as maps or curves, but is limited to bar-charts (including 100% bars and circles, or pie diagrams) and area bars.

In the geographic distribution there are more salient points to be noted. First, we should note the whole (such as world, continent, country, state, etc.) and the parts (continents, countries, state-groups, states, counties, cities, etc.), into which the whole is divided or distributed. Moreover, for convenience we should also note the number of parts, and their completeness (that is, whether or not their sum forms the total). Population and sales statistics are often complete, building statistics, and morbidity and mortality reports are examples of commonly incomplete data (compiled from only a few states or cities). In the graphic presentation of geographic distributions we can use maps in addition to the bar-charts which are applicable to all types of statistics, but we cannot use curves. For complete data we can generally shade or color whole areas on the map, but for incomplete data, such as reports for various cities, we should use isolated points on the map.

In the frequency series we need not only to know the basis of the numerical relations of the series and the number of items in the series, but we need also to know something very much

akin to completeness, namely its continuity. By continuity is meant whether the independent variable be a discrete or a continuous series, and if continuous, whether it be "point" or "period" data. By point data we mean data for separated non-contiguous points in the range; by period data we mean data for connected contiguous or overlapping periods in the range. For all these forms we can use curves, as you know, in addition to the ubiquitous bar-charts. But for the discrete series, generally composed of integral varieties, we are limited to the staircase or rectilinear curve so closely akin to bar-charts. Indeed this type of distribution has often little more than the accident of numerical designations to distinguish it from the abstract distribution. For the continuous series, generally of graduated variates, we should give a truer picture by using the smoothed curve or frequency polygon, though for period data we make a sacrifice of the accuracy of area representations thereby. And we may note a further distinction that while for discrete and period-data continuous series we should plot the data in the spaces between ordinates, for point-data continuous series we should plot the data on the ordinates. These distinctions have been discussed in the chapter on amount-of-change frequency curves.

Lastly we come to the historical series. Here four salient features of the independent variable should be noted. There is the range of time (the whole) covered, the intervals of time (or parts) used, and hence the number of items, and finally their regularity and continuity. The range and intervals are in centuries, decades, years, months, weeks, days, hours, and so forth. And some tables use different intervals in the same distribution, such as a series of decades followed by the individual recent years and lastly the most recent months, making in all for irregular intervals. The continuity of the data here means simply a point and period distinction. Items for isolated points of time, such as stock or balance reports at the first of each month, or price quotations at the end of each week, are point to point data. These the economists call stock or fund figures. Totals for periods of time, such as production or shipment statistics, are period data. These the economists call stream or flow figures. And it is of course possible to have isolated periods reported without the intervening periods. All historical data can be presented on curves, as well as bars, but in the summary chart the period data is

shown by bars, the point data by curves. In general, wherever we wish such a refinement, we can perhaps more accurately show period data by staircase curves or vertical bars, and point data by smoothed curves. A far more frequent distinction, however, is that while period data can be more accurately plotted in spaces between ordinates, point data can only be plotted upon the ordinates. These considerations have been brought out in various portions of the text.

To the foregoing discussion of statistics as regards the independent variable, we have now to add a brief outline of what may happen to the dependent variable. And here we can no further escape another general distinction which can be made between what might be called primary and secondary or derived statistics. The primary statistics consist of totals or averages which represent the original observations reported in the statistical table. Now these totals or averages, which, by the way, form the dependent or adjectival variable in the table, can be subjected to statistical treatment and materially modified, and the statistics which have suffered such treatment may be called secondary or derived data.

There are two main kinds or processes of statistical treatment which can take place simultaneously or individually. For the sake of simplicity they may be called compilation and conversion. Statistical compilation is to some extent operative upon both the independent and the dependent variables. Statistical conversion is almost entirely limited to the dependent variable and though it is perhaps considerably the more intricate subject, will receive scant attention from us, as it does not greatly affect the charting method.

Statistical compilation begins with material in its crude form, which is a mere listing or list. In the case of the logical distributions this is also about where it ends. Much rearrangement is possible, of course, but the abstract and geographical distributions always remain nothing more than lists. This, perhaps, is why neither can be shown in any graphic form where connection-lines represent continuity or sequence, in short, in any form of curve. The numerical distributions, however, can be so arranged that the items follow each other seriatim in a sensible way. In the early chapters of the book we have spoken of this as a case where the stubs or independent variable facts fall into an order imposed by themselves.

Obviously it is nothing more than mere mathematical sequence in the stubs, which dictates this order.

When a numerical distribution has been arranged in this orderly way it ceases to be a crude list and becomes a specialized one which is commonly called a "series." It is now usually ready for charting in the form of a curve, for there has appeared in the data a thread of connection running through the various items, a thread which enables us to connect the items on the chart by a line. This curved line can be shown on scales of all the various types discussed in the sections on curve-charts, from simple arithmetic or amount-of-change scales to logarithmic, rate-of-change, and other projections. It is not to be thought that the curve is something radically different from the bar-chart; indeed, as you know, the amount-of-change curve is simply a convenient sort of short-hand symbol of a series of bars, and the rate-of-change curves are merely special warpings and distortions of the amount-of-change curve with intent to bring out hidden relations and features in the curves. But it remains true that in the curve chart we are really for the first time, able to shift the focus of our attention from the comparisons or changes between individual items in a distribution, to something more complicated, the comparison or changes between these changes in different parts of the same distribution or in the same parts of different distributions. And to make the curve, we must first compile the numerical list into a numerical series.

This first step in statistical compilation is important, because to the casual reader it is hardly apparent. Indeed, it may be that the layman, glancing over a volume of the census, or over any other statistical series, is quite often under the innocent impression that the figures he sees just grew, somewhat as did Topsy. He does not suspect that many days or months of study may have gone into the determination of the proper group or interval limits in the series, and that over a year thereafter whole batteries of clerks and computing machines have sorted and enumerated the items in accordance thereto. Not all compilations are the result of such great attention. Regrettable it is to say, that publications still occur in which the raw material, the crude list, is given; but the compilation of data into series has not been carried far enough. Such lists and imperfect series tabulations are very likely to puzzle the student unless he detects the unfinished

treatment and completes the process of orderly series arrangement.

The simple series is the major step in the tabulation of numerical data but it is by no means the last, if the data be of the "period" type. If the data refer to scattered, isolated points of time in a historical series or points in a continuous frequency series, it may not be feasible to subject it to the processes about to be described. But period data (including discrete frequency series) in which the periods covered by the items are co-terminous and contiguous, can be subjected to the familiar process of cumulation and moving total (and average) calculation. In effect these processes change the separate groups or intervals into overlapping groups. In the cumulation, the overlapping is in one direction only; in the moving total (or average) (taken in its proper position as at the middle of its period) the overlapping is in both directions.

Slight differences may be observed in the susceptibility of the two numerical distributions to these processes of overlapping. Thus the frequency series can be cumulated in either direction, either backward or forward, yielding a "more than" or a "less than" cumulative. It may then be plotted in the familiar form of the ogive curve, of which both axes may be either arithmetically or logarithmically projected and for which the probabilities curve is the great analytical medium. The historical series, on the other hand, is sensibly cumulated only in one direction, and the curve of the cumulative is most especially used in the Zee-chart and its bars in the Gantt Progress chart. The historical series, moreover, can be subjected to moving total and average calculations, for any length of time or periodicity, and the moving series is to be used in curve charts of all kinds, while the frequency series is never subjected to this process except in some statistically technical calculations of the mode and smoothing processes. The subject of moving totals and averages and cumulations has been discussed in detail in the text.[1]

We come, lastly, to the other form of statistical treatment by which secondary data can be derived from primary statistical sources. It relates normally to the dependent variable and is the process of conversion of absolute data into relative

[1] Beyond the cumulative and moving or progressive totals, which are somewhat in the nature of integrals within limits, the processes of differentiation and integration are not included in this discussion.

data of various kinds. The absolute data is the original data itself, whether statistically compiled or not. It occurs in the form of totals or averages, which measure items in terms of various units of volume or value. The relative data is always the result of comparing this absolute data with other absolute data. The latter, with which comparison is made, may in general be called the base of the relative data. The relative data itself is of several varieties, which we will briefly mention, and occurs under many different names. Its relativity is usually obvious enough but is sometimes so completely disguised by ambiguous titles or nomenclature as to be difficult of detection and when this occurs its analysis may prove a baffling problem to the inexperienced.

The most familiar form of relative data is the percentage, or series of percentages, of which the total is 100%. It is the result of comparing the parts to the whole, or the items in a list or table to the sum thereof. In particular we may note that the base is common and constant, the same base being used for each and every percentage. Hence the percentages bear the same relation to each other as the original numbers or quantities which they represent bear to each other. The percentages, in fact, are but the same statistical facts reported in terms of a new unit of measurement. And so the percentage series may be subjected to cumulative and moving total (and average) calculation almost as readily as the original numerical quantities.

The next important type of relative data is that to which the special name "relative figures" is given. These share with the percentage figures the feature of a common and constant base; but differ from them in the relation to their base. Relative figures are not to their base-figure as parts to a whole; their sum does not total one hundred per cent. The relation of relative figures to their base is an item to item relation; that is, it is the result of comparing the original numbers, sometimes called by distinction, the "numerical data," with one of the component items. When the base-figure is more or less imaginary, being a combination of several often incommensurable base-figures, the relative figures are called "index numbers," though the latter term is by some writers loosely applied to all relative figures. Owing to the use of common and constant bases (for each series), relative figures and index numbers can

be subjected to moving total and average calculation, but their cumulation is usually of no value and significance.[2]

The last group of relative data differs from the foregoing in the use of various and different bases within each series instead of a common and constant base. Such series are always formed by comparing the items in one series with corresponding items in another series. Hence the items in the second series are the bases for the items in the resulting relative data. Where the latter stand in the relation to their bases of parts to wholes, they form percentages. Where they are not in this relation, the most common result is a per capita figure, "per family," "per dealer," etc. In either case when the fraction of ratio is very small it is usual to multiply by a thousand or some other constant, and so achieve a "rate." These relatives come in a wide variety of ways and under an equally wide variety of names. But they all have in common the feature of shifting inconstant bases. And as a result they cannot ordinarily be cumulated or smoothed by the moving total (or average) process. When we desire to compile them in these ways it is only proper to return to the original data and the base-figures and perform the operations upon them, that the smoothed relatives may be secured from their comparison.

At this point we draw to a close a very rapid survey of types and varieties of statistics as such. It is our hope that the reader will find such a panoramic view of value in his statistical work and the charting problems that arise therefrom. It is our hope that in time the classification and cataloging of statistical forms will become so simplified and improved that it can be used with immediate profit by the novice, as floral keys guide the amateur botanist to the name and description of the wayside flower. There is no reason why this codification and systematizing cannot take place, the structure of forms is really very simple, and the writer has no patience with the pernicious, though often unconscious, attempt to throw dust in the eyes of the layman and make technical problems appear more difficult than they really are.

He who has read with a broad comprehension the text to which this survey of statistical forms is the conclusion, will

---

[2] A minor variation of both this and the next type of relative, is the chain-percentage or link-relative, in which the base is not constant, but is always the preceding figure in the same series. It is a form of differential or successive differences. See rote in Chapter XXVI, page 307.

understand that graphic forms as well, are varieties and variations of a common root illustration. He will know that this common root picture is the representation of a single number by a straight line. He will know that any collection of numbers, be it an abstract, geographical, or numerical distribution, can be presented graphically by a collection of straight lines, which, if joined end to end, form a 100% bar; but if placed side by side form a bar-chart. The development of the pie-chart from the 100% bar, he will understand as merely a substitution of the circular for the straight line. The development of the curve he will understand as merely the connection and epitomizing of the bar-chart, suitable only for numerical series, whether frequency or historical. The varieties and modifications of curves will have no mystery for him. The area and three-dimension charts will stand before him as amplifications and combinations of bar-charts and curve-charts, suitable for interlocking composite distributions. The map will be but a variant of the latter, in which two dimensions of the paper picture the independent variable in geographic distributions, by longitudes and latitudes. The 100% triangle, the nomograph, and the calculating charts will be but patterns in which the two dimensions of the paper are devoted to the laborious illustration and proof of the very simplest propositions in geometry. This is really all there is to charts.

CHAPTER LVIII

THE FUNCTION OF CHARTS

A world turning to a saner and richer civilization will be a world turning to charts. From this conclusion, unwelcome as it may be, there is no escape. The case for the chart may even be sketched in a few schoolboy syllogisms, woven through the related ideas; civilization, clean-cut thinking; precision of thought, numerical statements; statistics, charts. With the last step in this chain, this book has attempted to deal. With a brief summary of statistical data the last chapter has provided us. There is no need to dwell upon the importance of precise, clear thinking, either in business or in economic studies. It remains to glance ahead a bit at the mechanics of the relations which charts will assume with the civilized world at large, and to venture a few predictions as to the nature of these relations.

And for this larger view it seems well to begin by amplifying our original definition. A chart is an image or graphic representation of abstract relations. Where these relations are not of a numerical nature, the chart is non-mathematical in character and is closely akin to the other graphic arts of a purely pictorial character; indeed its only distinction from paintings, photographs, and the like, appears to lie in the abstract nature of the ideas which it diagrammatically or schematically expresses. But where the relations are numerical and the subject of the chart is statistical, the chart is mathematical in character and forms a distinctly new branch of the graphic arts. While the artist will seek to present two groups of ten and twenty horses each by a picture of so many horses, placing his emphasis upon the realistic likeness of his drawing to horses, the chart-maker will seek to present the same objects by, let us say, two bars, which by their lengths express the numbers twenty and ten. His chart of horses will be exactly like his chart of two similar groups of ships, or his

chart of two very much larger groups of horses in which the group proportions are unchanged. He can, indeed, with equal facility make a chart for groups of two million and one million horses, a task which would be beyond the powers of the artist.

In subject-matter, then, the chart is universal, and hence, too, in its potential appeal and usefulness. No one can think of two numbers and attempt to comprehend their significance without, at least unconsciously, visualizing them; the number which does not conjure up in our minds some picture of quantity remains meaningless to us.

In this sense, therefore, everyone who deals with numbers is already a chart-maker and a chart-user. We have no choice between the use of charts and the use of statistics; we have only a choice between the use of written or physical charts and the use of imagined or "mind's-eye" charts. Often, indeed, the latter are sufficient, and many persons, it is true, still prefer under all circumstances to carry all the pictures of their numerical data in their minds. But for the careful study of important figures, or for the casual study of large bodies of important figures, this is obviously the less efficient method, and the physical record, the written or graphic chart, comes into service. It is more permanent, more convenient, and more accurate.

The technique of the chart is also, in a sense, wider than that of the other graphic arts; indeed, it comprises something of the technique of all the arts. The reader of this book has seen that we have drawn statistics with pictures, and sculpted them as models and we have reproduced them by photography and by lantern slides and by printing. In this we have freely used design, relief and color. It may not be too much to add that some day we shall set charts to music, to enhance their graphic value, evolving a musical expression of statistics. This will seem less improbable when we consider its use in the accompaniment of moving pictures of charts.

The animated chart, made possible by the motion-picture film, has long been a dream of the author. Its graphic value will be great in the presentation of fundamental economic facts to the general public, or of special statistics to special audiences. By its means the important chart can be presented in various stages of completion, and attention can be focused in turn upon each change, development, or addition to the picture. Thus in a bar-chart, the labels can appear first, then each bar,

with its data, can appear, one after the other, until the bar-chart is completed. Curves can be shown wiggling across co-ordinate rulings, with close-ups of each important added wiggle. Maps can appear first in outline and the shadings can appear and spread across the map by simple tricks of photography, and these shadings can be altered to show changing conditions for successive points or periods of time. The "movie" of statistics is clearly coming, for schools and colleges, for the general public, for the scientific or academic meeting, and in business, for director's meetings, for sales conventions, and for advertising purposes.

In all chart-making, a distinction which will become increasingly recognized is the distinction between charts for popular consumption and charts for research purposes. This is no more than adapting the chart to the audience for which it is intended. And there can be as many different proper charting ways as there are different degrees of familiarity with charts and ease in chart-reading. For extremely popular presentation, the pie-chart is always effective; bar-charts should be converted into series of circles and curves into vertical bars whenever possible. For more sophisticated readers the amount-of-change curve can be used; for the technical and semi-technical, the simpler forms of the rate-of-change curve are permissible. The probabilities and other special projections will be really understood only by the experts; and are essentially charts for internal consumption in the research laboratory.

Though everyone can be told how a bow is carried across a violin string, we do not expect all to play the violin well. And though the technique of chart-making can be simply explained, we cannot expect everyone to make good charts. The chief source of good charts will always be the statistical departments of large organizations. When the organization is an institution for the promotion of research in some special field, the statistical staff will of course be well manned. But the greatest strides, at least in chart-making, if not also in statistical methods, will in the future be made in the statistical departments of large business organizations.

In business the function of the statistician is two-fold, comprising on the one hand special research and investigations, and on the other hand, the co-ordination and intelligent reporting of current business operations. In both of these, charts

are essential implements. In the research field the statistician has often a scouting function, his job being to look ahead and try to forecast the future development of the house and its markets. In the reporting field he assembles and interprets the operations of all the other departments; purchasing, production, shipment, warehousing, sales, and other collections; and of the business as a whole: inventories, costs, and profits. His position here is that of liason or intelligence officer between the responsible head of the business and his subordinates, and also between the responsible subordinates and their departments. To get the fullest use of the expert intelligence in visualization and analysis, one of the vice-presidents may be himself a professional statistician and chart-maker of the highest specialized training, but in the past the average statistician has not often displayed a sufficiently practical viewpoint to justify this connection and the wealth of significance which lies in the records of the individual business house is untapped by those who must guide it.

Comparable to the lawyer who brings to the guidance of business enterprise an intimate knowledge of legal technicalities, is the business statistician who brings to it an intimate knowledge of statistical interpretation. In business houses where the operations and problems are of a standardized nature, his skill will not, except in very large concerns, be constantly needed and the statistician here becomes a consulting expert rather than a permanent officer. In such concerns the reporting procedure can be quickly set in motion and standardized, so that it can be carried on thereafter by clerks. The Gantt progress-charts and a few of the simpler curves and maps are all that need be installed, after the proper system of records from the accounting and other departments have been established. In business concerns of more variety of operations, the trained statistician is necessarily more of a permanent member of the personnel and the work of forecasting is likely to be seriously entered into. Here the widest variety of charts come into use, for a nice understanding of their graphic value and true significance is available. Here it is often profitable to maintain a special statistical "laboratory" with complete facilities for statistical sources, compilation and analysis and for graphic records.

The well-furnished statistical department should, of course, contain the mechanical calculators, the double-entry adding

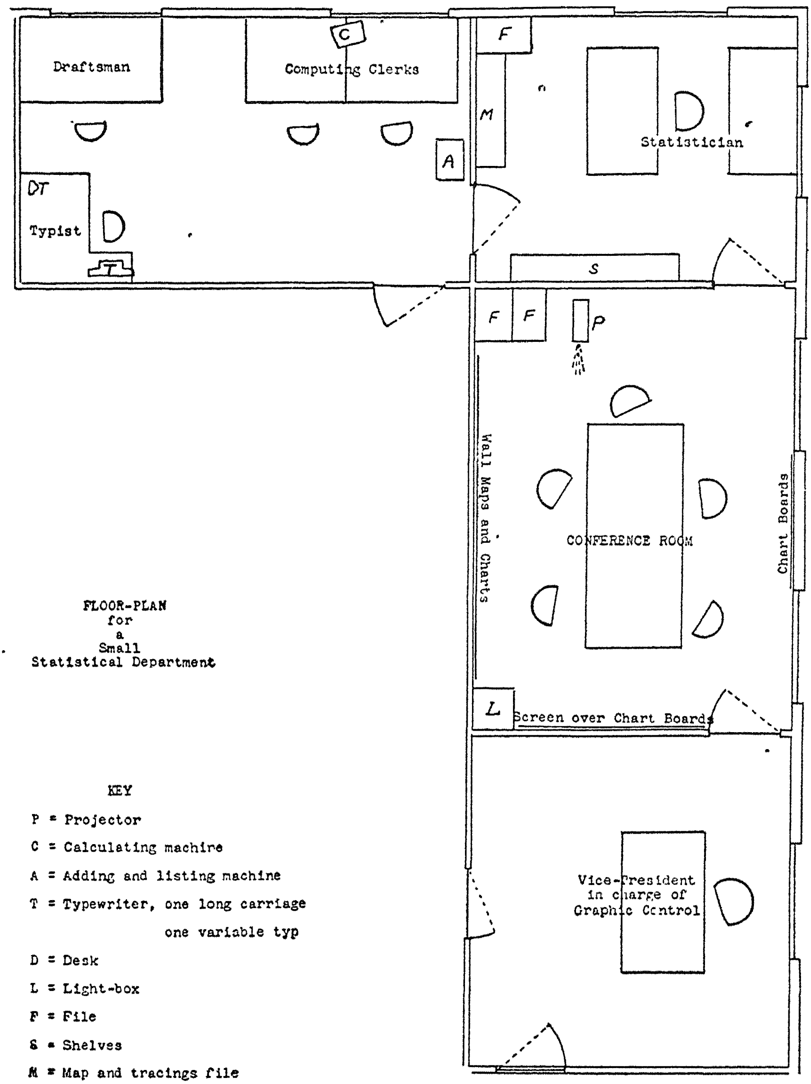


Fig. 496.

and listing machines, the multiplying and dividing machines, and perhaps the card-punching multiple-entry machines (the true posting machines). It should contain full drafting facilities and perhaps also blue-printing or photostating machinery. These things obviously belong to the workroom, which should be apart from the office of the directing and

creative statistical officer. But the department is not complete without full accommodations for the successful study of its results. There should be a conference room, convenient to and properly fitted for the use which will be made of it by the directors or vice-presidents and responsible heads of the business.

In the conference room the charts perform their chief function in business, as guides to the formation of policies. The room should be equipped with a light-box for the comparison of curves and with a screen and projector for charts which it is desirable to exhibit to several persons at once. Important data can be permanently posted on large wall-boards and these wall-boards, by the use of sliding panels, can hold large bead-maps as well. All important data should be on record in chart form, either in looseleaf binders or vertical files. Needless to say, the room should always be locked up when not in use and the keys to it should be in the hands of but two or three persons. It should be the repository for all information about the concern which is of value in the formation of policies, this information being in chart form because of the ease with which it can then be consulted.

Of a much more general nature are charts for popular consumption. These are appearing with increased frequency in newspapers, general magazines, and technical publications. The day will come when no statistical compilation will be regarded as complete until it is illustrated with charts which present its major significance. The greatest development of charts, here, however, will take place in the advertising columns, and in general for propaganda work. For the proper chart is an excellent weapon against the inertia, indifference, and often hostile attitude of the average reader. It is not merely the best kind of eye-catcher for calling attention to numerical data, it is also the most convincing proof of that data. The most casual reader stops a moment before any diagrammatic puzzle to examine it. If he finds incidentally that he immediately understands it, he is perhaps at once pleased with it and is sure at least to carry away with him a memory of the message it conveyed. That such charts should be of the simplest, goes without saying; and here too, expert skill in chart-making is desirable. For the right chart is strong in inverse ratio to the technical ability of the reader,

and the less effective would be text or tables, the more powerful grows the right chart.

In all fields, scientific, academic, and commercial, the chart is a medium of expression too forceful to be overlooked, too valuable to be neglected. Its future growth will assuredly be rapid and perhaps in many ways even startling. In this book we have set forth many ways for the presentation of statistics and statistical relations. The category is, however, by no means complete. It cannot be complete, for the charts are still in the making, and the methodology of graphic illustration is in no sense that of a perfected art. There is room for much improvement in existing chart-forms as well as in the development of altogether novel forms. New ideas will come out of the research laboratories, new methods, new forms, new charts. The distinction between graphs for popular publication to the general public and graphs for internal consumption in the statistical workshop, will become more marked; and as public knowledge increases, charts will pass out of the workshop into the magazine and book page, no less through advertising than through text columns.

We are finding a new language, the grammar of which is not yet completed, nor the dictionary written. It is well that this is so, for codification and systemization easily bring stagnation; and volumes such as the present, in which the existing material is set in order, must not be allowed to stifle new growth. The reader is urged not to permit the rules laid down in this book to restrict his efforts, but rather to allow the principles set forth to stimulate his imagination and enterprise. The pictorial display of mathematical and numerical statements is an illustrative art, with the high object of facilitating human understanding and vision, an end the achievement of which justifies all means, be they orthodox and accepted or novel and previously untried.

FINIS

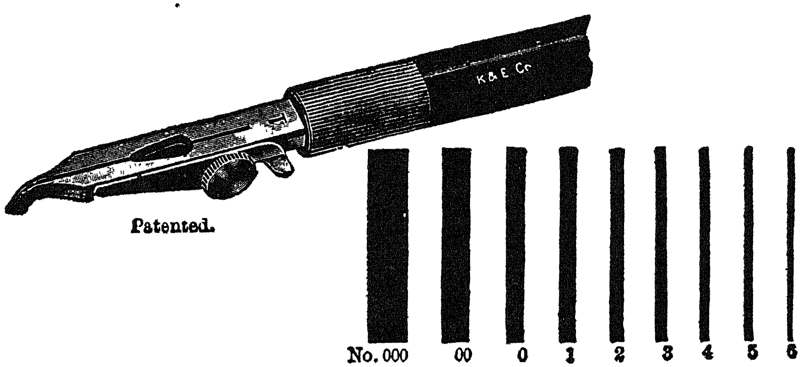
APPENDICES

APPENDIX A

IMPLEMENTS FOR MAKING CHARTS

The equipment which should be available in a chart-making or statistical office depends, of course, a great deal upon the types and forms of charts which will be developed and the nature of the data which will be handled statistically. Charts can be made at home or in the very smallest office, with nothing more than a draftsman's ruling pen, some India ink, and good paper. As the chart-making work grows, more drawing pens will be added in order that different colored inks can be ruled in without delays, and in order that several operators can be drawing at the same time. A good drawing board is necessary and perhaps two drawing boards are best, one about 24 by 30 inches for large charts, and the other about 18 by 24 inches for smaller charts. Together with the drawing board, there should be a T-square and small triangles or tri-squares. Occasional need arises for one or two patterns of French curves. A good drawing set includes a compass or dividers for the drawing or circles of circular outlines. Several dotting machines for the drawing of dotted, broken, and dotted dash lines are on the market, and when they can be successfully used, save considerable time, but they are not always successfully used. A section-liner is almost essential for much cross-hatching work. A protractor is necessary where circles will be divided into proportional parts and angles must be used. For bar-charts, a double ruling pen, or railroad pen, is a great convenience, adding to the appearance of the chart by making more absolute the uniformity of the bar widths, and greatly decreasing the amount of time required for the making of the bar-chart. (This pen rules in two parallel lines simultaneously.) Bar-charts can also be made on large scales with adhesive tape or passe-partout; and several mechanical bar-charts are now on the market in which cloth tape is unwound from invisible or hidden spools and drawn out to the

required length of the bar-chart. These last are useful where the length of the bar-chart must be frequently changed and brought up to date (as in a sales-manager's office where the bars represent the weekly averages or cumulatives of the work of the individual salesmen). In map work, a curved ruling pen is sometimes an advantage for the drawing of rounded curves; better still, is the Payzant lettering pen. A planimeter is



Permission of Keuffel & Esser, N. Y.

Fig. 497. The Payzant Lettering Pen.

often useful for checking up total areas on the map. A pantograph is a device by which outlines can be copied on larger or smaller scales, and is sometimes useful in map work; very cheap pantographs can be obtained, which will ordinarily be satisfactory.

A straight-edge is useful for the cutting of paper, and the best knives available are the ordinary one-sided razor blades, as the paper-cutting knives require frequent sharpening. The straight-edge is necessary because from time to time in the cutting and trimming of paper the edge itself will be cut into by the knife and if the T-square has been used the T-square will then be ruined. Ordinary camel's-hair brushes are often used for applying water-color or ink to maps; but the best way of coloring maps, as has been previously described, is by the use of wax crayons, the cheaper and waxier the better, the wax being afterward removed by a sharp knife-edge or razor blade, leaving the desired color tint. For the quick filling in of bars and solid areas, the lettering pen is desirable. It can be obtained in many sizes, but the results are not entirely even and smooth unless the operator is skilled. A special pen (the Payzant) has been put out for fine lettering,

which has a round nub with a well similar to that in a drawing-pen, holding considerable ink. This pen is also available in several sizes.

Typewriters should be used as far as possible in the lettering of charts, as well as in the entry of data upon the charts. The process is much more rapid than hand-lettering, and the results are, for the average-sized chart, usually better. There are two sizes of standard typewriting type, "pica" (10 characters horizontally to the inch) and "elite" (12 characters horizontally to the inch). Unless large type is especially desired to facilitate small reproductions, the elite is better, for it enters all figures in less space. Any special type-faces, such as Gothic or Italic, may be had, but the usual type-face is Roman. The figures come in two styles in all machines, "book-keeping" and "regular." Book-keeping type has slightly greater visibility but gives uneven lines, as the numerals have swinging tails. The regular numerals are usually more satisfactory.

The typewriter carriage to use for chart-making should accommodate 11-inch paper (and larger, if charts are being made on sheets larger than $8\frac{1}{2}$x11 inches). One standard machine (the Royal) will take 11-inch paper on its regular carriage, all other makes require long carriages. The typewriter to use for chart-making should also so hold the paper as to print down to the very bottom edge of the paper without shifting. There is but one standard machine (the Royal) which will do this. Besides the standard machines, the Hammond has good features for chart-making in that it will print any style of type at a moment's notice, 9 lines to the inch instead of 6, and will space the characters properly (and if desired, as close together horizontally as 18 characters to the inch); but this machine is not so easily handled by most operators as the standard machines, because it has a three-shift key-board, and, while it will accommodate any size of paper, the paper is likely to shift slightly in it and cannot be used down to the bottom edge.

For the statistical work, special computing machinery is desirable, both for its speed and for its accuracy, being for these reasons, where the amount of statistical work to be done is considerable, a great economy. Computing machinery is in general of two types, the adding machine and the calculating or multiplying machine. The adding machine can be obtained

in both printing (or "listing") and non-printing types. Obviously, the printing machine is far more desirable because the operation can be checked back by an examination of the printed page or record. A special type of adding machine, called the duplex model (Burroughs), is desirable for work in which totals of parts will be required. These part totals are known as transfer-totals, the machine having two faces, the one of which clears without removing the record from the other face, so that several adding operations can be conducted at the same time. The duplex machines are particularly useful in the compiling of statistics by States, the part totals being taken off for State groups, and afford a great saving in the detection of errors when the data is checked over. Still a third type of adding machine is the mechanical tabulator (such as the Hollerith), which works with punched cards in which the amounts to be entered with full descriptive detail, are represented by holes in a card, and these holes operate the machine, just as do the holes in the records of the player-piano and similar devices. This is the true posting machine—it sorts, posts, and adds, with a typewritten record if desired, automatically.

The calculating or multiplying machines so far manufactured are only of the non-print type and leave no written record. This, of course makes errors more difficult to detect. There are two general types of machines. The first (such as the Burroughs or Comptometer) automatically adds as fast as the operator punches the keyboard, and calls for specially trained operators who can punch several keys at once and continue punching the same keys the proper number of times to effect the multiplication (shifting the position of their fingers on the keyboard for each new digit in the multiplier). The second type of machine performs the same operation automatically from a single punching of the keyboard. The operator may be required to turn a handle the proper number of times to effect a multiplication, and to shift the recording dials one space for each digit, but the work is very quick and absolutely accurate. In the latest model German machines, electrically driven, the work is entirely automatic after the setting of the keyboard. Adding and calculating machines are economical of time and expense if electrically driven.

All of these calculating machines are virtually multiple-adding machines, for effective quick addition. They are

accurate to the last figure recorded. For many statistical purposes, however, such as the figuring of percentages, accuracy is not necessary beyond the third or fourth figure. For such work, a slide-rule is sufficient and, though harder on the eyes, is much more portable and is indispensable in the statistical office. The accuracy of reading increases with the length of the rule; and slide-rules are made five, ten, and twenty inches long. Further accuracy can be gained by the use of magnifying glasses fitted on the runners of the rule. Needless to say, slide-rules will perform many operations outside of the powers of the calculating machines.

Appendix B

STEPS IN MAKING CHARTS

In statistical offices, both large and small, it is desirable to have as near as possible an approach to what may be called "straight-line methods" of chart production. Only by the institution of such methods can the routine work of a large number of charts be satisfactorily and economically accomplished. For this purpose, it is desirable to break up the work of chart-making into various steps and stages, and to have the individual charts, as they pass through these stages, pass from the desk of one operator to another in a regular series and direction. The various steps outlined below will be found to be a fairly complete list of the stages through which different kinds of work will pass. Very often, however, some one or more of these stages will be omitted for particular charts and for particular data.

I. The first step in chart-making is, of course, the collection of data or statistics, namely, the information to be shown upon the chart. This information can be gathered in two different ways. The first way ought, wherever possible, to be followed out, whether the second is used or not.

The first way to gather data is to consult and collect all information previously compiled by other investigators on the particular subject. This is a class of research work ordinarily involving the consultation of the various books, pamphlets, and other authorities in the public libraries, and calling for the services of fairly skilled library workers. A reasonably complete knowledge of the various sources and authorities in which the particular information sought is likely to be found in its most useful and complete form, is of course desirable, in order that the search will not take too much time. When the information has been found, it can be carefully copied upon specially prepared work-sheets or data sheets by the investigator, or, if the data is in compact form, it can

be photographed or photostated and the photostatic copies used in the office. The latter method, though apparently more expensive, is far more accurate and reliable and more economical of time, so that, in the end, it is generally the cheaper process.

The second method of gathering the information is by the use of field investigations. It is an independent and original study for the purpose of securing primary information rather than secondary information (i.e., information taken second-hand from other investigators). The field investigation may be carried out by a skilled investigator, if it is not too extensive and if the resources are available for sending a thoroughly skilled investigator out. When this is not possible, or when the extent of the investigation is very great, the method of questionnaires can be used. In this case, the art of questionnaire-making comes into play, for the drawing up of a questionnaire is by no means as simple as it might appear.

A good questionnaire is one in which no question can be misunderstood; one in which each question is capable of only one meaning and of a precise answer of one type only. Moreover, a person who has drawn up a questionnaire must be able, beforehand, to envisage his entire problem, foreseeing all moot points and issues which will arise in the course of the investigation, and to which answers will be desired. Lastly, the questions must be so framed as to avoid, as far as possible, any psychological reactions either upon the part of the investigator or of those whom he questions and consults for his information. In fact, the psychological difficulties about many questionnaire problems are the principal obstacles, and the method of questionnaires is fast losing ground for precise statistical compilation, because of the careful analysis and psychological interpretation, translation, and correction to which the answers must be subjected before they can be satisfactorily compiled. And great as is the task of preparing and conducting a satisfactory questionnaire, the problem of compiling its answers is sometimes even greater, and requires a staff of more than ordinary intelligence.

And in both research and field investigation work, wherever clerical tasks have been performed, it is, of course, necessary that careful checking be done on all such clerical work in order to catch and correct errors which are humanly inevitable. A definite place in the schedule of preparing charts must be

given to this work of checking back for accuracy on all the clerical work performed.

II. The second step in the routine of chart-making is ordinarily called the computing. This often requires that the data be first copied upon specially prepared forms or work-sheets, so that it may be subjected to the proper processes. The computing should, as far as possible, be planned considerably in advance in order that work may progress evenly and smoothly. The work is of a clerical or statistical nature and calls for the services of statistical or computing clerks, and frequently also for a battery of adding and calculating machines. The work-sheets which are to be used for the job should be carefully designed with an eye to the machinery by which the processes of calculation can be most easily performed. Thus, in cases where totals and subtotals are desired, the work-sheets should be designed so as to fit into the adding-and-listing machines in order that the machine may operate directly on the work-sheet and not upon the usual tape. If the work is first done on the tape, it will have to be copied on the work-sheet, with the unavoidable percentage of error and the great additional time and labor involved. The only alternative is to paste the tape on the work-sheet, and this will not be possible if the work-sheet is not large enough. Needless to say, for all computing steps the most rigorous checking for errors is necessary. If possible, the computing should be so done that it is self-checking, or easily checked for accuracy by a single checking operation performed upon the totals for the data. It is best, therefore, when the adding machines are to be used, to have work sheets in which the lines coincide with the adding machine lines, so that the sheet can be run through the machine instead of tape. Moreover, if possible, the items to be added should be entered in a column, with blank lines between them, so that the machine entries may appear immediately below the hand-written entries, this reduces error and facilitates checking.

III. The next step is the beginning of the chart-making. The general character of the charts to be used should have already been determined. Suitable chart-forms should have been obtained from some publisher, or made to order by one's own printer. These forms should accommodate the data in typewriting. And this step, therefore, may be called "entering up the chart." It calls for the services of an intelligent and

capable "tabulating typist," for the work must be both accurately and neatly done. If, in the case, for example, of curve-charts, the ordinates at which the data must be entered have been placed at uniform typewriter intervals of one-third of one inch, the typist will be able to work rapidly and smoothly, and work will cross this desk promptly. One good typist, under such circumstances, can keep half a dozen compiling and drafting clerks busy, and will generally find time to do computing work as well. Again, checking for errors is necessary.

IV. When the final data has been entered upon the chart, the next step is one of plotting this data in chart form. The immediate proximity of the data to be charted upon the chart itself, as placed there by the last operation, makes this drafting or plotting process extremely easy, and where the chart-fields have been already printed or drawn upon the paper, the drafting requires little more than the proper selection of plotting points and the careful ruling in of curved lines or bars. The work calls for the services of an ordinarily intelligent clerk and only in the case of very complicated charts, or in cases where extra lettering and entering of data will be done by hand, is it necessary that skilled draftsmen be employed. The draftsman should work under the best available light, as the eye-strain of careful plotting or ruling is severe. Art school students and engineering school students are generally qualified to perform this work capably. Again, the process of checking must be carefully done to detect error.

V. The final step on any chart is the labelling and the finishing up of the various details left unfinished in the typewriting and plotting stage. Unless the form of title for the chart has been standardized, the problem of a correct, complete, and easily understood title for a chart is sometimes very difficult. And the title should, in general, be made by the department head or statistician who is responsible for the work. There should always be a portion of the chart sheet in which the title can be conveniently located where it will be at once apparent to the reader. In the case of historical curve-charts, where the field is low on the page, the title naturally belongs at the top. In other cases where the chart is higher up on the page, the title can be placed below the chart.

A good title should not merely give the nature of the phenomena shown by the chart together with the general kind of analysis followed by the chart, but it should also give such

distinctive details as will separate the chart from all others in the series. Where the series of charts is clearly connected, as in historical curves, the date or year of the individual chart can be placed in the corner of the paper, the chart title being reproduced alike on all charts in the series. In addition to the title of the chart, the chart should also tell its source, that is, the authority for the information it presents. And, in addition to these two items which must be typed upon the page, there is generally considerable labelling of data columns. Sometimes there is labelling of individual curves upon the chart-field itself.

When all this has been done and the chart has been dressed up in its final form, it should be carefully inspected by some one competent, as far as possible, to detect errors which appear in the chart; and the chart should have in one corner a place for the "O. K." signature of the person in authority who has finally approved of the chart. In addition to this approval signature, the data of approval should be shown so that charts of different date of manufacture can be easily seen and the latest and most reliable chart distinguished.

Personnel.—In short, it will be seen that the statistical office has, in the main, five major processes, namely: research work, computing, typing and printing, drafting, and inspection (including titling), and, in general, it may be said that these processes call for different types of workers, namely: research workers or librarian, statistical clerks, typists or letterers, draftsmen and artists, and statisticians.

Notes.—In addition to the finished chart itself, it is sometimes desirable to have appended notes or explanatory comments which will serve to interpret the significance of the chart to the reader or executive who will consult it, and which will point out to him the important facts displayed by the chart. The writing and composing of these explanatory notes should be composed by the statistician in charge. These explanatory notes should be in the most easily comprehensible form. Their composition calls for an extremely practical understanding of the point of view of those who will read and consult the chart, and requires a return to the language and to the non-technical line of thought which will be pursued by the layman.

Reports.—Nothing has been said here about the mobilizing and assembling of a large number of charts upon one subject

into the form of a single coherent report. This is usually considered a matter for the skill and judgment of the statistician himself. According, however, to so excellent an authority as Mr. Charles P. Steinmetz,[1] there are three kinds of reports, and the most complete report generally contains these three types within itself. The first is the general report in which the final conclusions and significances are summarized briefly, the report perhaps taking up about 10 per cent of the entire report. It is this part of the report, and often this part only, which will be read by many executives, or the average reader. In the second part of the report, these conclusions which appeared in the summary or general report are expanded in greater detail to show their bases or foundations and to enable the careful reader to delve deeper into individual phases and aspects. This second part may contain about 30 per cent of the total number of pages in the report. The third part of the report is the technical authority and technical detail which will be read only by those who are extremely anxious to check up upon the work of the compiler, either for the sake of repeating or elaborating the investigation, or for the sake of detecting errors or confirming the accuracy of the information given. This part may often take up the greater portion of the report.

[1] Steinmetz, Charles P., *Engineering Mathematics*, McGraw-Hill Book Co, 1917, p. 290-293.

APPENDIX C

METHODS OF PRESENTING CHARTS

In the main, the charts described in this book have been discussed upon the basis of presentation upon ordinary size letter paper, that is, paper measuring 8½x11 inches. Occasionally, larger sheets of double this size have been mentioned, and in the section on models the need for mounts and containers of uniform size has been discussed. The 8½x11-inch paper is perhaps the most generally convenient because of its conforming to standard sizes of office paper and vertical and other filing methods. In an office where legal size paper is used, it would obviously be better to adopt sheets 8½x13 inches as the standard chart size, and on occasion, to use sheets of double this dimension.

In the section on curves, which form the great majority of charts, the desirability of positioning the curve in one corner of the paper for ready comparison and of leaving large margins above and to the left of the chart-fields has been explained. The chart-form itself should be carefully designed as the one most suitable to the type of chart-work which will be done.

It is of the greatest importance that the "field," or rulings, of the chart-form be in a faint ink, preferably green or gray. (The orange and reds are hard on the eyes; the blues will not photograph.) When the charts are to be reproduced on a much smaller scale, it is well to make all but the more important co-ordinates in blue, so that they will not be reproduced. For this purpose, blue co-ordinate paper can be used, the co-ordinates which should be reproduced being ruled in by hand in black ink.

The usual types of maps are not ordinarily printed upon standard sizes of paper, and if a great deal of map-work is to be done, it is well to have special maps printed upon regular sizes of paper to conform to the rest of the charts in use. Maps

on the $8\frac{1}{2}$x11-inch paper, often of inferior quality, can be obtained from a few manufacturers of charting materials. These are sometimes better than more elaborate maps, as they are generally only outline maps showing State or county boundaries.

Attempts have been made to establish standard charting forms upon cards for card-catalogue filing, small curve-charting fields being printed upon 4x6-inch cards in one corner of the card. The fields are ruled off for amount-of-change curves only, with ordinates for 52 weeks, 12 months, and 31 days, in the usual way for historical curves. The amount of data which can be entered upon these card charts is, of course, limited and the detail in which the curves are shown is not great. The thickness of the card prevents, to a certain extent, the facility of "light analysis." Some publishers present these small charts on thin paper suitable for tracing or "light analysis," as part of a loose-leaf note-book system to be carried about in one's pocket.

A major problem in graphic presentation arises when the charts are to be shown to large audiences. The practice is often followed of making the drawings extremely large, say three by four feet in size, on heavy paper which can be unrolled and pinned against a bulletin board for display. These rolls of paper, however, are difficult to carry about and are easily damaged. This is even more true when heavy card-board is used, which can not be rolled but must be carried about flat. The best advice appears to be to present the chart upon tracing cloth, which can be fastened at one end upon a large stick of wood and easily rolled or unrolled. Where expense is no consideration, the window-shade rollers on which school maps are mounted can be used. These are contained in long narrow boxes to keep the chart dust-proof, the chart being unrolled by pulling the lower end out of the box in the same way that a window shade is drawn down.

Much the best method for the display of charts of large size to an audience is by the use of lantern slides and a lantern slide projecting machine. These machines can be purchased for small sums in a very handy shape, folding up like valises and easily portable. The lantern slide can be made by any photographer at small cost, from the original chart used in the office made in the usual way upon ordinary size paper. Where colored areas are shown upon the chart, the same colors

may be shown on the lantern slide, by coloring the slide with Japanese transparent photographic colors. This coloring work will be done by the photographer according to instruction, or according to the original chart which has been photographed. Lantern slides are smaller than post-cards and easily carried about. Except through breakage, they are not easily damaged.

The lantern slide method has more to recommend it in the fact that a drawing which will be seen upon the lantern slide will easily be visible to the entire audience. Where the very large original drawings are used instead of lantern slides, the lines in the drawings have to be made very much thicker to make them visible, but with lantern slides, the faintest lines, if distinctly visible upon the plate, will be clearly projected to the entire audience. A safe rule is that whenever the original office copy of the chart, from which the lantern slide is made, is larger than 8½x11 inches, the lines upon it will not be clear and definite upon the lantern slide (because of the reduction in size) unless the lines are made heavier; but in the case of originals up to 8½x11 inches, not only will the ordinary markings be clear and distinct but the ordinary typewritten labels and data will also be visible.

A very expensive machine has been devised, which is useful in large offices for the display of many charts to board meetings and other small audiences. It is a projecting machine which does not require lantern slides, but which will reflect the image of the original chart upon the screen. The convenience of this type of machine is, of course, very great, as it eliminates the delays and expense of lantern slides, and permits the exhibition of any material at a moment's notice, even though the need for such an exhibit had not been foreseen. The machine is, therefore, valuable where a large number of charts may be shown and it is not certain beforehand which ones will be desired.

The method par excellence for popular audiences is the moving picture film and machine, with the charts shown as actually developing and building up. The manufacture of such films is not easy and is very expensive, but the results fully justify it. Thus, a bar-chart shown in this way would first appear merely as a list of items, and then, one by one, the bars would appear upon the screen, until the entire chart was assembled. Such a chart receives careful study and all its parts

are understood, and their significance is grasped by the audience.

The reproduction of charts in large numbers involves special problems. Where only a few copies are desired, photostating is the best method. Blue-printing is a more economical method but its results are neither so clear nor so attractive. By the use of blue-printing, only negatives can be obtained, in which black areas appear as white and white areas appear as blue. A somewhat similar method is known as the Van Dyke process, or black-line and brown-line process. These are obtained partly by offset and partly by photographic methods, and positives as well as negatives can be secured, but the original chart should be upon very translucent, un-water-marked paper (best of all, upon tracing cloth) and very distinctly and clearly drawn. In these two methods, the process is a photographic "printing-through" one, the light passing through the chart to a sensitized surface. It is necessary, therefore, that no extraneous matter be upon the reverse side of the chart and that no corrections be made upon the chart by overlaying fresh paper or Chinese white. It is also desirable that the chart paper be clear, un-water-marked, translucent. In these respects, the blue or sepia print, and the black-line or brown-line print are unique. As they are "printing-through processes," it is always advantageous, when there is typewriting upon the original drawing, to back up the sheet with reversed carbon paper, so as to get an additional and coinciding imprint of the typewriting upon the reverse side of the original.

The above limitations and precautions do not apply to the truly photographic processes, that is, the photostat, the photograph, and the photo-engraving. In these, the light is reflected back from the surface of the chart to a sensitized surface, without passing through the chart. Thick, opaque paper can be used, with corrections in Chinese white or on special slips of paper pasted over the incorrect parts; also, the condition of the back of the chart does not matter. When only a few copies are needed, photostating is the best process, far more convenient and only slightly more expensive than blue-printing. The least expensive course is to get a photostatic negative and as many blue-print positives as desired. When sufficient copies are desired to make printing advisable (that is, printing from metal plate) it is necessary to use either the "line-cut" or the "half-tone."

The line-cut is the more economical but shows only full black and white markings. The half-tone (in which the object is photographed through a screen) is a more expensive process, giving results with all variations and tints of grey as well as almost full white and full black. Half-tones are much more expensive than line-cuts and require more time in their manufacture, but the results, of course, are better when areas are shaded with different tints and colors in the original chart. The line-cut is sufficient for the ordinary bar-chart or curve-chart and can often be given elaborate shadings and cross-hatchings by the use of the Ben Day process.

For all photographic methods, including printing, care must be taken in the choice of colors, when colors are used upon the original chart, for as has been previously pointed out the various colors reproduce differently by photography than would be expected from their appearance to the eye, and certain shades which appear decidedly different to the eye, may be exactly similar to the camera and reproduce alike. Red, of course, photographs as black (that is, appears as black upon the photographic print), while blue does not photograph at all but appears as white upon the photographic print, and yellow appears almost black. Thus it will be seen that a scale from red to blue arranged in chromatic sequence of colors will, to a certain extent, photograph as a natural sequence of grays ranging from solid black to white in the photographic print.

Three other methods of reproduction are in ordinary commercial usage, the hectograph, the mimeograph, and the multigraph. The first two of these can be used satisfactorily for the reproduction of charts. In the hectograph, the old-fashioned jelly-offset process, special hectograph ink can be used in several colors which will simultaneously reproduce, the copies appearing somewhat paler than the original but reproducing color for color at a single offset. The number of copies that can be secured from hectographic offset is, however, limited. The best offset processes claim to afford as many as 50 to 70 copies from a single original, but as a rule 25 to 30 are all that will be satisfactory. When not more than two or three dozen copies are desired, the hectographic process is extremely simple and satisfactory in the average office, its only requirement being that the special hectographic or copying ink or pencil, or typewriter ribbon be used in the making of the original chart which is to be copied.

The mimeographic process is essentially a stencilling one, the chart being first drawn upon a fine wax or fibre stencil and then laid over the drum of an inking and printing machine, the ink passing through the cuts in the stencil and printing upon paper. The mimeographic process will produce as many copies as desired up to several hundred, but is limited, like ordinary printing, to one color only, for the color is determined by the ink which has been previously placed upon the drum of the printing machine. The mimeographic process can be conveniently used where mimeographic reproductions are already being made of typewritten copy. Its limitations are that it shows only one color, and that it is impossible to reproduce solid black areas on the mimeograph. Shading must be done by cross-hatching, and extensive cross-hatching is apt seriously to damage the mimeograph stencil. Moreover, the mimeographic process may prove a dirty one, even for those who are experienced with it and certainly for the beginner. And it is extremely difficult to adjust the stencils upon the printing drum so accurately and so evenly that the reproduction will be precise; as a rule, a slight curvature or wrinkling of the stencil is observable; this, of course, makes for irregular, curved, or broken lines on the chart when straight lines are desired.

There is a special instrument known as the mimeoscope which is well adapted to the drawing of charts upon mimeograph stencils. It is also useful in the office as a "light-box" for general light analysis. The machine consists of a strong electric light under a ground glass (ground to diffuse the light). By the use of this machine, charts can be easily traced on mimeograph stencils, using the various kinds of mimeograph styles provided for the purpose. The mimeograph method is adapted to the reproduction of simple charts in one color only, without solid areas, and not requiring precise reproduction, when the number of copies desired is between 50 and 500; for a smaller number of copies, the hectograph processes are satisfactory, and for a larger number of copies the printing processes are more economical.

In a few cases where charts are largely prepared upon the typewriter, and not more than two, three, or four copies are desired, the carbon copy method of reproduction can be used. Carbon paper can be obtained in red as well as black, and sometimes in other colors, and the first two or three carbon copies

are likely to be very good when the paper on which the charts are made is not too thick or soft. Sometimes the labelling and typing for a number of copies can be done at one operation on the typewriter with carbon paper, or in two operations when different colors of carbon paper must be used (first printing up all of one color and then substituting the other color of carbon paper and printing up that). Drafting upon the charts must be done on the various copies independently with drawing pens in the usual way, making each of the copies original so far as the drafting is concerned unless, as in the case of bar-charts, the chart can be made with the typewriter. Bound carbon sheets (in binders) help to prevent shifting of the copies.

Whenever typewriting is done for charts which are to be blue-printed, or photographically copied by direct printing (in which case, the light will have to pass through the entire chart-paper), it is best to insert a piece of carbon paper behind the original chart and print a reverse copy upon the back of the paper to add intensity to the typewriting on the chart. Only fully inked typewriter ribbon should be used for charts which are to be photographically copied, as it is desirable that the printing should be as intense as possible for the best photographic results.

Appendix D

COLORS IN CHARTS

A great deal of controversy takes place over the use of colors in charts. In a previous appendix we have discussed those limitations to the use of colors which arise from the needs of specified reproduction processes. In this appendix will be considered the use of colors in cases where no mechanical limitations are imposed and colors can be judged entirely upon their own merits.

The argument against colors takes the line that the average business man is not accustomed to them, and will consider more carefully a product in the familiar black and white. The colors in this view tend to make the chart "pretty" and prettiness is rightly to be avoided. Other things being equal, artistic effects are always desirable, but never to the point where they attract attention and distract the reader's mind from the message of the chart.

But even in this view, not all lines need be equally strong and it is freely conceded that the co-ordinates of the field of a chart should be as light as possible, and, even better than in black, ruled or printed in gray ink. For the field is only the background, and the lighter shade throws the curve or plotting more distinctly into the foreground. And the best practice has already gone further and given to the field of a chart invariably the color of green. The green should be medium light with somewhat more yellow than blue, so that it will photograph as well as gray. The advantage of the green is that no confusion with black plotting or lettering is possible. The rule may be extended to make it universal, even the maps to be used in statistical reports being printed in green outlines.

Red is a color to which the accountant is accustomed to give a negative significance, and may well be used for contrast with black in curves and map-shadings, wherever it

applies to opposite data. Thus on a map the red should be used for unfavorable conditions, and in curves for costs or expenses, and the like. It is one of the great advantages of the red that the corresponding data (and scales, if any) can be typewritten in the same color without difficulty when the typewriter is equipped with a bi-color ribbon. Brown, green, and blue typewriter ribbons can also be secured, but require special shifting of ribbons in the machine. Red and black are the main colors used in the charting office.

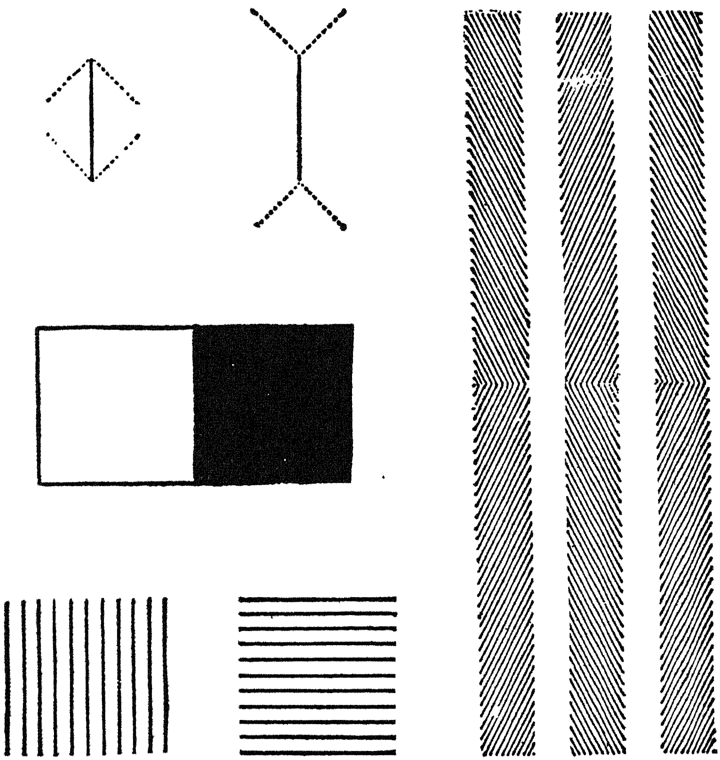
Blue is taboo in chart-work and should rarely be used. Its least fault is that it is hard on the eyes when used steadily in the place of black. Its great fault is that it does not photograph, and will therefore disappear if the chart is photostated or blue-printed or photographed in other ways. In maps it may be used for shading to indicate favorable conditions, with the understanding that when photographed, these areas will be white; in curve work it may be used for lines, or rulings which are intended to disappear when photographed, either for secrecy or to eliminate details useful only in plotting and undesirable in reduced copies. When blue is used to disappear either in ink or typewriting, care should be taken to see that it contains no red pigment, as this will defeat the purpose. When a great deal of red is present, we have, of course, purple—another color which should not be used, as it looks black under most artificial light.

Yellow and green, are little-used colors, but sometimes serve in curve-work as secondaries to black and red, as for example where it is desired to insert "quotas" or comparisons in fainter colors, on the same chart with black and red curves. These colors are also useful in map work, when areas are shaded, as intermediates between the red and blue extremes—the best sequence being red, orange, yellow, yellow-green, and blue-green. The other colors, brown, pink, and the like, are very little used, as they are not so distinctive as the simple primary colors.

Appendix E

OPTICAL ILLUSIONS IN CHARTS

Mention has been made of the danger in bar-charts, and area-charts, of making one area appear larger than another merely through the use of more powerful shading. This



By permission of the publishers of the "Book of Knowledge," 2 West 45th Street, N. Y. City.

Fig. 498. Optical Illusions.

applies to colors as well as to the use of various grays and cross-hatchings. But there are no hard-and-fast rules as to what

is a powerful shading, as each case depends upon the surrounding colors or shades with which it is in contrast. The chart-maker must in each case judge carefully of the effects of his shadings, and even if he cannot give equal emphasis to all parts of his chart, at least strive to avoid emphasizing the unimportant and slighting the important parts of its message.

The accompanying illustrations show that, in the particular case presented, white is more powerful than black and the white square appears larger than the black although the black has a border or outline added to it. They also show that a line may be made to look longer or shorter by the direction of arrows attached to it, that hatchings make the same rectangle look wider or narrower according to their direction, and also that bars hatched diagonally may be made to appear crooked instead of straight. These are but a few of the minor optical illusions against which the maker of the chart should be on guard.